

Sr. Big Data/Hadoop Developer Sr. Big Data/Hadoop Developer Sr. Big Data/Hadoop Developer -
Aetna Inc Hartford, CT Over 9+ years of IT experience in software development and support with
experience in developing strategic methods for deploying Big Data technologies to efficiently solve
Big Data processing requirement. Proficient in installing, configuring and using Apache Hadoop
ecosystems such as MapReduce, Hive, Pig, Flume, Yarn, HBase, Sqoop, Spark, Storm, Kafka,
Oozie, and Zookeeper. Strong comprehension of Hadoop daemons and MapReduce topics.
Strong knowledge of Spark for handling large data processing in streaming process along with
Scala. Hands On experience on developing UDF, DATA Frames and SQL Queries in Spark SQL.
Experience in integrating Kafka with Spark streaming for high speed data processing. Ability to
develop MapReduce program using Java and Python. Good understanding and exposure to
Python programming. Exporting and importing data to and from Oracle using SQL developer for
analysis. Extensive use of Open Source Software and Web/Application Servers like Eclipse IDE
and Apache Tomcat. Experience in designing a component using UML Design-Use Case, Class,
Sequence, and Development, Component diagrams for the requirements. Experience in
understanding the security requirements for Hadoop and integrate with Kerberos authentication and
authorization infrastructure. Strong experience in various phases of Software Development Life
Cycle (SDLC) as requirement gathering, modeling, analysis, architecture design, development,
testing and implementation. Strong experience on designing Big data pipelines such as Data
Ingestion, Data Processing (Transformations, enrichment and aggregations) and Reporting.
Strong experience in submitting Spark applications in different clusters such as Spark Standalone
and Hadoop Yarn. Good knowledge on various Amazon Web Services (AWS) such as S3, EC2,
VPC, RDS, SQS, ELB. Experience in tuning and improving the performance of spark jobs by
exploring various options. Strong Experience in migrating data using Sqoop from HDFS to
Relational Database Systems and vice-versa. Strong experience in developing the workflows
using Apache Oozie framework to automate tasks. Experience in working with MapReduce
programs using Apache Hadoop to analyze large data sets efficiently. Strong experience in
working with Core Hadoop components like HDFS, Yarn and MapReduce. Strong experience in

Cloudera Hadoop distribution with Cloudera manager. Experience in launching spark applications by using the Kerberos authentication. Solid front-end developer experience in developing UI applications using JSP, Ajax, CSS, HTML, Java Script, XML. Hands on Experience in developing and implementing the Spring Rest and Restful Web Services. Strong knowledge in JAVA Messaging Service (JMS). Experience in generating logging by Log4j to identify the errors in production test environment. Efficient in developing java applications in various Integrated Development Environment (IDE) tools like Eclipse, My Eclipse and RAD. Excellent knowledge on Hadoop Architecture and ecosystems such as HDFS, MapReduce, Job Tracker, Task Tracker, NameNode, DataNode and Secondary NameNode concepts. Excellent knowledge in building and scheduling Big Data workflows with the help of Oozie and Auto-sys. Experience in Elastic Search (used for Faster Indexing), Kibana (Creating Dashboards), Splunk (Log Analysis and Dashboards).

Excellent working experience in Scrum/Agile framework and Waterfall project execution methodologies. Good Knowledge in Amazon AWS concepts like EMR and EC2 web services which provides fast and efficient processing. Experience in using Web Services Technologies: Web Services, RESTFUL, SOAP UI. Proficient in using and deploying applications to Web Servers/Application servers like Tomcat, WebSphere, Micro services. Working experience with distributed systems using JMS, Message Queues, Spring JMS Integration Hands on Experience in integration with Maven, JUnit and Log4j frameworks. Experience in writing Build Scripts using Shell Scripts, MAVEN and using CI (Continuous Integration) tools like Continuum, Jenkins. Work Experience Sr. Big Data/Hadoop Developer Aetna Inc - Hartford, CT May 2017 to Present Responsibilities Worked as a Sr. Big Data/Hadoop Developer with Hadoop Ecosystems components like HBase, Sqoop, Zookeeper, Oozie, Hive and Pig with Cloudera Hadoop distribution.

Involved in Agile development methodology active member in scrum meetings. Worked in Azure environment for development and deployment of Custom Hadoop Applications. Designed and implemented scalable Cloud Data and Analytical architecture solutions for various public and private cloud platforms using Azure. Involved in start to end process of Hadoop jobs that used various technologies such as Sqoop, PIG, Hive, MapReduce, Spark and Shells scripts (for scheduling of few

jobs). Implemented various Azure platforms such as Azure SQL Database, Azure SQL Data Warehouse, Azure Analysis Services, HDInsight, Azure Data Lake and Data Factory. Developed SQL scripts using Spark for handling different data sets and verifying the performance over MapReduce jobs. Used J2EE design patterns like Factory pattern & Singleton Pattern. Involved in converting MapReduce programs into Spark transformations using Spark RDD's using Scala and Python. Extracted and loaded data into Data Lake environment (MS Azure) by using Sqoop which was accessed by business users. Manage and support of enterprise Data Warehouse operation, big data advanced predictive application development using Cloudera & Hortonworks HDP. Developed PIG scripts to transform the raw data into intelligent data as specified by business users.

Utilized Apache Spark with Python to develop and execute Big Data Analytics and Machine learning applications, executed machine learning use cases under Spark ML and MLlib. Installed Hadoop, Map Reduce, HDFS, Azure to develop multiple MapReduce jobs in PIG and Hive for data cleaning and pre-processing. Used Spark API over Hortonworks Hadoop YARN to perform analytics on data in Hive. Improved the performance and optimization of the existing algorithms in Hadoop using Spark Context, Spark-SQL, Data Frame, Pair RDD's, Spark YARN. Developed Spark code using Scala and Spark-SQL/Streaming for faster testing and processing of data. Developed a Spark job in Java which indexes data into Elastic Search from external Hive tables which are in HDFS. Explored MLlib algorithms in Spark to understand the possible Machine Learning functionalities that can be used for our use case. Used windows Azure SQL reporting services to create reports with tables, charts and maps. Executed Hive queries on Parquet tables stored in Hive to perform data analysis to meet the business requirements. Used Oozie and Zookeeper operational services for coordinating cluster and scheduling workflows. Configured Spark streaming to receive real time data from Kafka and store the stream data to HDFS using Scala. Continuous coordination with QA team, production support team and deployment. Performed transformations, cleaning and filtering on imported data using Hive, Map Reduce, and loaded final data into HDFS. 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. Import the data from different sources like HDFS/HBase into Spark RDD and developed a data pipeline using Kafka and Storm to store data into HDFS. Used Spark streaming to receive real time data from the Kafka and store the stream data to HDFS using Scala and NoSQL databases such as HBase and Cassandra. Documented the requirements including the available code which should be implemented using Spark, Hive, HDFS, HBase and Elastic Search. Performed transformations like event joins, filter boot traffic and some pre-aggregations using Pig. Developed code in Java which creates mapping in Elastic Search even before data is indexed into. Configured Oozie workflow to run multiple Hive and Pig jobs which run independently with time and data availability. Imported and exported the analyzed data to the relational databases using Sqoop for visualization and to generate reports for the BI team. Support Cloud Strategy team to integrate analytical capabilities into an overall cloud architecture and business case development. Developed customized Hive UDFs and UDAFs in Java, JDBC connectivity with hive development and execution of Pig scripts and Pig UDF's. Used Hadoop YARN to perform analytics on data in Hive. Developed and maintained batch data flow using HiveQL and Unix scripting. Developed and execute data pipeline testing processes and validate business rules and policies. Built code for real time data ingestion using Java, MapR-Streams (Kafka) and STORM. Environment: Hadoop 3.0, HBase, Sqoop, Zookeeper, Oozie, Hive 1.2, Pig 0.17, Agile, Azure, MapReduce, Spark 2.3, J2EE, Java, Zookeeper 3.4, Oozie, Cassandra 3.11, NoSQL Big Data/Hadoop Developer HCSC - Chicago, IL July 2015 to April 2017 Responsibilities Extensively worked on Hadoop eco-systems including Hive, MongoDB, Zookeeper, Spark Streaming with MapR distribution. Developed Big Data solutions focused on pattern matching and predictive modeling. Worked on analyzing Hadoop cluster and different big data analytic tools including Pig, HBase database and Sqoop. Involved in Agile methodologies, daily scrum meetings, spring planning. Created RDD's and applied data filters in Spark and created Cassandra tables and Hive tables for user access. Worked on NoSQL support enterprise production and loading data into HBase using Impala and Sqoop. Performed multiple MapReduce jobs in Pig and Hive for data cleaning and pre-processing. Build Hadoop solutions for big data problems using MR1 and MR2 in YARN. Handled importing

of data from various data sources, performed transformations using Hive, Pig, and loaded data into HDFS. Involved in identifying job dependencies to design workflow for Oozie & YARN resource management. Performed Hadoop installation, configuration of multiple nodes in AWS-EC2 using Hortonworks platform. Worked with NoSQL databases like HBase in creating HBase tables to load large sets of semi-structured data coming from various sources. Involved in Hadoop cluster administration and successful in maintenance of large volumes of storage. Upgraded the Hadoop Cluster from CDH3 to CDH4, setting up High Availability Cluster and integrating Hive with existing applications. Designed & Developed a Flattened View (Merge and Flattened dataset) de-normalizing several Datasets in Hive/HDFS which consists of key attributes consumed by Business and other down streams. Analyzed the existing data flow to the warehouses and taking the similar approach to migrate the data into HDFS. Involved in PL/SQL query optimization to reduce the overall run time of stored procedures. Exported data from HDFS to RDBMS via Sqoop for Business Intelligence, visualization and user report generation. Involved in designing schema, writing CQL's and loading data using Cassandra. Built the automated build and deployment framework using Jenkins, Maven. Implemented MapReduce jobs in HIVE by querying the available data. Configured Hive Meta store with MySQL, which stores the metadata for Hive tables. Performance tuning of Hive queries, MapReduce programs for different applications. Proactively involved in ongoing maintenance, support and improvements in Hadoop cluster. Worked with cloud provisioning team on a capacity planning and sizing of the nodes (Master and Slave) for an AWS EMR Cluster. Worked on data using Sqoop from HDFS to Relational Database Systems and vice-versa. Maintaining and troubleshooting Exploring with Spark to improve the performance and optimization of the existing algorithms in Hadoop using Spark context, Spark-SQL, Data Frame, pair RDD's. Created Hive Tables, loaded claims data from Oracle using Sqoop and loaded the processed data into target database. Used Cloudera Manager for installation and management of Hadoop Cluster. Worked on MongoDB, HBase databases which differ from classic relational databases Involved in converting HiveQL into Spark transformations using Spark RDD and through Scala programming. Integrated Kafka-Spark streaming for high efficiency

throughput and reliability Worked in tuning Hive & Pig to improve performance and solved performance issues in both scripts. Environment: Hadoop 3.0, Hive 1.2, MongoDB, Zookeeper 3.4, Pig 0.17, HBase, Sqoop, Agile, NoSQL, Impala 3.0, MapReduce, YARN, Oozie, AWS, Hortonworks, Kafka, Cassandra 3.11 Sr. Java/Hadoop Developer T-Mobile - Bellevue, WA November 2013 to June 2015 Responsibilities Worked on analyzing Hadoop cluster and different big data analytic tools including Pig, Hive and Sqoop. Worked on implementation and maintenance of Cloudera Hadoop cluster. Assisted in upgrading, configuration and maintenance of various Hadoop infrastructures like Pig, Hive, and HBase. Developed presentation layer components comprising of JSP, Servlets and JavaBeans using the Spring framework. Responsible for developing various modules, front-end and back-end components using several design patterns based on client's business requirements. Designed and Developed application modules using spring and Hibernate frameworks. Involved in developing Use case diagrams, Class diagrams, Sequence diagrams and process flow diagrams for the modules using UML and Rational Rose. Developed the presentation layer using JSP, AJAX, HTML, Bootstrap, XHTML, AngularJS, CSS and client validations using JavaScript. Designed and developed SOAP & RESTFUL web services. Developed Java classes confirming J2EE design patterns, JNDI, packaged with J2EE specifications Developed and executed custom MapReduce programs, Pig Latin scripts and HQL queries. Used Hadoop FS scripts for HDFS (Hadoop File System) data loading and manipulation. Developed simple to complex MapReduce streaming jobs using Java language for processing and validating the data. Implemented Spark using Python and Spark SQL for faster processing of data. Developed Spark jobs and Hive Jobs to summarize and transform data. Involved in converting Hive/SQL queries into Spark transformations using Spark data frames, Scala and Python. Consumed Web Services using WSDL, SOAP and UDDI from third party for authorizing payments to/from customers. Developed and implemented the MVC Architectural Pattern using Spring Framework. Extensively used Spring JDBC in data access layer to access and update information in the database. Developed a RESTful API that provided account management capability as well as security role lookup and management for all downstream dependencies. Responsible for the design and

development of the application framework. Used spring's test framework to create integration tests for various spring boot and spring batch applications. Implemented data access using Hibernate (ORM Tool) persistence framework. Environment: Pig 0.14, Hive 1.0, Sqoop, HBase, Spring 4.5, JavaBeans, Hibernate 4.8, JavaScript, AJAX, HTML, Bootstrap, AngularJS, CSS, J2EE, Java, HDFS, Spark, Python, Scala, MVC Sr. Java/Full Stack Developer CoreLogic - Austin, TX April 2012 to October 2013 Responsibilities Involved in various phases of Software Development Life Cycle (SDLC) of the application like Requirement gathering, Design, Analysis and Code development. Implemented MVC architecture and DAO design pattern for maximum abstraction of the application and code reusability. Developed web application using JSP custom tag libraries, Spring Action classes and Action. Designed Java Servlets and Objects using J2EE standards. Configured the application using Spring framework annotations and developed Spring Controllers for request and response processing and implemented RESTful Web Service. Migrated Spring based application to cloud based Micro services. Designed and developed the REST based Micro services using the Spring Boot, Spring Data with JPA. Participated in coding Spring AOP components for the Transactional Model to handle many requests. Involved in developing Java APIs, which communicates with the JavaBeans. Implemented Java/J2EE Design patterns like Business Delegate and Data Transfer Object (DTO), Data Access Object. Used JSP for presentation layer, developed high performance object/relational persistence and query service for entire application utilizing Hibernate. Developed views using Bootstrap components, Angular-UI and involved in configuring routing for various modules using angular UI router. Extensively used HTML, JavaScript, Angular.js and Ajax for client side development and validations. Developed Web services (SOAP) through WSDL in Apache Axis to interact with other components. Created Stateless Session EJB's for retrieving data and Entity Beans for maintaining User Profile. Used Log4j as logging framework to capture the log traces of applications in debugging the issues. Used ANT automated build scripts to compile and package the application. Designed database and created tables, written the complex SQL Queries and stored procedures as per the requirements. Involved in Unit, Integration and Performance Testing for the new enhancements.

Developed the application using Java Beans, Servlets and EJB's. Environment: SDLC, MVC, JSP, J2EE, Java, JavaBeans, Hibernate 3.5, Bootstrap, HTML, JavaScript, Angular.js, Ajax, Log4j, ANT

Java Developer Impetus October 2009 to March 2012 Responsibilities As a Developer in Java, involved in System Requirements analysis and conceptual design. Experienced in Web development with JavaScript, JQuery, HTML, Bootstrap, CSS, and Ajax. Implemented Java batch jobs for nightly runs and worked heavily on concurrency API for a low latency high throughput application. Worked on applications, included in requirement analysis, Design, Development, and Testing. Used ANT, Maven to build & deploy applications, helped to deployment for CI using Jenkins and Maven. Designed and developed Application based on Spring Framework, Spring MVC and Spring templates. Created JUnit test cases for unit testing the code at minute level and used Eclipse IDE. Developed and designed new cross-browser accessible interfaces using JQuery and JavaScript. Involved in writing JSPs, JavaScript and Servlets to generate dynamic web pages and web content. Developed Use Case Diagrams, Object Diagrams and Class Diagrams in UML using Rational Rose. Worked on Spring IoC, Spring MVC framework, Spring Messaging Framework and Spring AOP to develop application service components. Used Hibernate ORM framework for persistence to database by integrating it with Spring framework using Spring Hibernate template. Used JavaScript and JQuery for providing client-side validation and Spring Validators for server-side validation. Used JQuery to make the frontend components interact with the JavaScript functions to add dynamism to the web pages at the client side.

Developed User interface using JSP, Angular JS, JSP Tag libraries, third party libraries and JavaScript. Used JavaScript, JQuery and Ajax API for intensive user operations and client-side validations. Used Hibernate framework in persistence layer for mapping an object-oriented domain model to a relational database. Used Spring Core for concept Inversion of control (IOC) implemented using dependency injection. Developed Object Model and UML design models for developing Use cases and created Sequence diagram, class diagram and active diagrams for application components and interfaces. Defined and developed the application's presentation layer using HTML, CSS and JavaScript. Developed the User Interface using JSP and used CSS for

style setting of the Web Pages. Involved in developing applications for workflow using JSP's, spring MVC module, Hibernate, AJAX, JavaScript technologies using Apache Tomcat. Implemented design patterns like DAO, singleton, factory to achieve design principles Used Jenkins for continuous integration and Maven for building the EAR file. Generated JUnit test cases for testing various Java components. Environment: JavaScript, JQuery, HTML 4, Bootstrap, CSS, Ajax, Java, ANT, Maven, Jenkins, Hibernate, MVC, JUnit Education Bachelor's Skills AJAX (5 years), Apache (6 years), APACHE HADOOP HDFS (5 years), APACHE HADOOP MAPREDUCE (5 years), APACHE HADOOP SQOOP (5 years), APACHE HBASE (5 years), Bootstrap (5 years), Database (9 years), Hadoop (5 years), HADOOP DISTRIBUTED FILE SYSTEM (5 years), HBase (5 years), HDFS (5 years), Hive (5 years), HTML (5 years), Java (7 years), JavaScript. (5 years), JSP (5 years), MapReduce (5 years), Servlets (5 years), SQL (7 years)

Name: Keith Powers

Email: xgray@example.org

Phone: 210.665.0041x985