**Anne Bowen**

[annebowen@fairhint.com](mailto:annebowen@fairhint.com)

**2529660203**

**Summary:**

* Over8years of professional IT experience with 5+ plus years of experience in analysis, architectural design, prototyping, development, Integration and testing of applications using Java/J2EE Technologies and 3+ years of experience in Big Data Analytics as Hadoop Developer with good knowledge in Hadoop ecosystem technologies.
* Delivery experience on major Hadoop ecosystem Components such as Pig, Hive, Spark Kafka, Elastic Search &Hbase and monitoring with Cloudera Manager.Extensive working experience using Sqoop to import data into HDFS from RDBMS and vice-versa.
* In-depth experience and knowledge in developing and analyzing Map Reduce Jobs and Applications developed standalone and/or through Pig/Hive.
* Extensive experience in developing Pig Latin Scripts for transformations and using Hive Query Language for data analytics.In depth knowledge of Spark concepts and experience with Spark in Data Transformation and Processing.
* Hands on experience working on NoSQL databases including Hbase, Cassandra and its integration with Hadoop cluster.Experience in development and utilization of Apache SOLR with Data Computations and Transformation for use by Down Stream Online Applications.
* Good experience in ETL tool Informatica.
* Solid experience in developing job workflows and schedules with Oozie, and IBM Tivoli
* Experience in Hadoop administration activities such as installation and configuration of clusters using Apache, Cloudera and AWS.Working on Automation using Perl, Shell,Python and Scala Scripts.Developed UDFs in Java as and when necessary.
* Training and Knowledge in Mahout, Spark MLlib for use in data classification, regression analysis, recommendation engines and anomaly detection.Automating the jobs using Unix shell scripting and providing production support.
* Training and Knowledge in Mahout, Spark MLlib for use in data classification, regression analysis, recommendation engines and anomaly detection.
* Good experience in Python.Developed UDFs in Java as and when necessary to use in PIG and HIVE queries.
* Working knowledge of database such as Oracle 8i/9i/10g, Microsoft SQL Server, DB2, Netezza.
* Good experience in Oracle Business Intelligence Enterprise Edition(OBIEE)
* Experienced in using Version Control Tools like SubVersion, Git.Experience in development of logging standards and mechanism based on Log4J.Good understanding and experience with Software Development methodologies like Agile and Waterfall.
* Experienced in design, development, Unit testing,integration, debugging and implementation and production support, client interaction and understanding business application, business data flow and data relations from them.

**Technical Skills:**

|  |  |
| --- | --- |
| **Hadoop/Big Data Technologies** | HDFS, MapReduce, Spark, Hive, Pig, Sqoop, Flume, HBase, Cassandra, Oozie, Zookeeper, YARN, Hue, Ambari, Kafka, Elastic Search, Cloudera, Hortonworks, Tez, Apache Parquet |
| **Programming Languages** | Java JDK1.4/1.5/1.6 (JDK 5/JDK 6), C/C++, Unix Shell Scripting, Python |
| **Web /Applications Servers** | Apache Tomcat, Weblogic, WebSphere and Bastion. |
| **Operating Systems** | UNIX, Windows, LINUX |
| **Databases** | HBase, Oracle 8i/9i/10g, Microsoft SQL Server 2008/2012, DB2 & MySQL 4.x/5.x, Teradata, Mongo DB |
| **Java IDE** | Eclipse 3.x, IBM Web Sphere Application Developer, IBM RAD 7.0 |
| **Tools** | TOAD, SQL Developer, SOAP UI, ANT, Maven |

**Professional Experience:**

**American Express,Houston, TX Oct 14 – Present**

**Sr. Big Data/Hadoop Developer**

**Responsibilities:**

* Coordinated with business customers to gather business requirements. And also interact with other technical peers to derive Technical requirements and delivered the BRD and TDD documents.
* Extensively involved in Design phase and delivered Design documents. Experience in Hadoop eco system with HDFS, HIVE, PIG, SQOOP and SPARK with SCALA.
* Worked on analyzing Hadoop cluster and different Big Data Components including Pig, Hive,Spark, HBase, Kafka, Elastic Search, database and SQOOP. Installed Hadoop, Map Reduce, HDFS, and developed multiple Map-Reduce jobs in PIG and Hive for data cleaning and pre-processing.
* Importing and exporting data into HDFS and Hive using SQOOP.
* Migration of 100+ TBs of data from different databases (i.e. Netezza, Oracle, SQL Server) to Hadoop.
* Written Hive jobs to parse the logs and structure them in tabular format to facilitate effective querying on the log data.
* Involved in creating Hive tables, loading with data and writing hive queries that will run internally in map reduce way. Generate OBIEE reports to verify the Hive tables data.
* Experienced in defining job flows.Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting. Experienced in managing and reviewing the Hadoop log files.
* Used Pig as ETL tool to do Transformations with joins and pre-aggregations before storing the data onto HDFS. Responsible to develop data pipelines from different sources
* Utilized Apache Hadoop environment by Cloud era Distribution.Exported data from HDFS environment into RDBMS using Sqoop for report generation and visualization purpose.
* Worked on Oozie workflow engine for job scheduling.Involved in Unit testing and delivered Unit test plans and results documents.

**Environment:** Hadoop, MapReduce, HDFS, Hive, Pig, Hue, Ganglia, Nagios, Java, Kafka, Elastic Search, SQL, Scala, Oracle, Netezza, Ambari, Sqoop, Flume, Oozie, Java (jdk 1.6), Eclipse.

**Kindred Healthcare,Arlington,TX Mar 12 – Sep14**

**Big Data Tools Expert / Hadoop Developer**

**Responsibilities:**

* Hands on experience in loading data from UNIX file system and Teradata to HDFS
* Experienced on loading and transforming of large sets of structured, semi structured and unstructured data from HBase through Sqoop and placed in HDFS for further processing.
* Installed and configured Flume, Hive, Pig, Sqoop and Oozie on the Hadoop cluster.
* Involved in creating Hive tables, loading data and running hive queries in those data.
* Extensive Working knowledge of partitioned table, UDFs, performance tuning, compression-related properties, thrift server in Hive.
* Involved in writing optimized Pig Script along with involved in developing and testing Pig Latin Scripts
* Working knowledge in writing Pig’s Load and Store functions.
* Developed Java MapReduce programs on log data to transform into structured way to find user location, age group, spending time.
* Developed optimal strategies for distributing the web log data over the cluster, importing and exporting the stored web log data into HDFS and Hive using Scoop.
* Collected and aggregated large amounts of web log data from different sources such as webservers, mobile and network devices using Apache Flume and stored the data into HDFS for analysis
* Monitored multiple Hadoop clusters environments using Ganglia.
* Developed PIG scripts for the analysis of semi structured data.
* Developed and involved in the industry specific UDF (user defined functions).
* Used Flume to collect, aggregate, and store the web log data from different sources like web servers, mobile and network devices and pushed to HDFS.
* Analyzed the web log data using the HiveQL to extract number of unique visitors per day, page views, visit duration, most purchased product on website.
* Integrated Oozie with the rest of the Hadoop stack supporting several types of Hadoop jobs out of the box (such as Map-Reduce, Pig, Hive, and Sqoop) as well as system specific jobs (such as Java programs and shell scripts).
* Monitored workload, job performance and capacity planning using Cloudera Manager.
* Managing and scheduling Jobs on a Hadoop cluster using Oozie.

**Environment:** Amazon EC2, Apache Hadoop 1.0.1, MapReduce, HDFS, CentOS 6.4, HBase, Kafka, Scala, Elastic Search, Hive, Pig, Oozie, Flume, Java (jdk 1.6), Eclipse, Sqoop, Ganglia, Hbase.

**Krogers Stores, Marietta,GA Oct 10 – Feb 12**

**Java Developer**

**Responsibilities:**

* Responsible for requirement gathering and analysis through interaction with end users.
* Involved in designing use-case diagrams, class diagram, interaction using UML model with Rational Rose.
* Designed and developed the application using various design patterns, such as session facade, business delegate and service locator.Developed UDFs in Java as and when necessary to use in PIG and HIVE queries.
* Worked on Maven build tool.Involved in developing JSP pages using Struts custom tags, JQuery and Tiles Framework.
* Used JavaScript to perform client side validations and Struts-Validator Framework for server-side validation.Developed Web applications with Rich Internet applications using Java applets, Silverlight, JavaFX.
* Involved in creating Database SQL and PL/SQL queries and stored Procedures.Implemented Singleton classes for property loading and static data from DB and Debugged and developed applications using Rational Application Developer (RAD).
* Developed a Web service to communicate with the database using SOAP.Developed DAO (data access objects) using Spring Framework 3 and Deployed the components in to WebSphere Application server 7.Actively involved in backend tuning SQL queries/DB script.Worked in writing commands using UNIX, Shell scripting.
* Involved in developing other subsystems’ server-side components.Production supporting using IBM clear quest for fixing bugs.
* Generated Java wrappers for web services using Apache AXIS

**Environment**: JBoss, XMLSOAP, RESTful, Java EE 6, IBM WebSphere Application Server 7, Apache-Struts 2.0, EJB 3, Spring 3.2, JSP 2.0, Web Services, JQuery 1.7, Servlet 3.0, Struts-Validator, Struts-Tiles, Tag Libraries, ANT 1.5, JDBC, JMS, Service Bus.

**GAP Inc,Montgomery, AL Mar08 – Sep10**

**Java Developer**

**Responsibilities:**

* Involved in Design, Development and Support phases of Software Development Life Cycle (SDLC). Used agile methodology and participated in Scrum meetings.
* Developed the application using Spring Framework that leverages Model View Layer (MVC) architecture UML diagrams like use cases, class diagrams, interaction diagrams (sequence and collaboration) and activity diagrams were used. Developed UDFs in Java as and when necessary.
* Data from UI layer sent through JMS to Middle layer and from there using MDB message retrieves Messages and will be sent to MQSeries.
* Used JSON as response type in REST services.
* Used RESTFUL client to interact with the services by providing the RESTFUL URL mapping.
* Performed unit testing using JUnit.Developed UDFs in Java as and when necessary to use in PIG and HIVE queries.
* Provided direction and support for technical architecture with respect to performance, business continuity and seamless integration/functioning of applications, databases, servers, networks.
* Managed, administered and maintain more than 100 Oracle databases.

**Environment:**Java SE 6, Servlets, XML, HTML, JavaScript, JSP, Hibernate, Oracle 11g, SQL Navigator.