

BIG DATA DEVELOPER

For Further Assistance please contact James today.

PROFILE

Hadoop Developer

2 years of experience in Big Data processing using Apache Hadoop

5 years of experience in development, data architecture and system design.

EDUCATION

Jawaharlal Nehru Technological University, India

Bachelor of Technology in Electronics and Communication Engineering

TECHNOLOGIES

Languages: C, C++, Java, JavaScript, HTML, CSS , VB.

Big Data Technologies: Apache Hadoop, Map-Reduce, HDFS, Pig, Hive, Hbase and Sqoop

Databases: Oracle 9i/10g/11.2.1, MS SQL 2000/2005/2008, MS Access.

Tools: MS PowerPoint, MS Word, MS Excel.

App/Web Servers: Tomcat 3.3/5.0/6.0 and JBoss. 4.0

Operating Systems: UNIX (Solaris, Linux), Windows NT/2000/XP/2003

Development tools: Eclipse, Visual Studio, MYSQL Workbench, SQL+

Architecture Tools: Microsoft Visio 2002/2003, Rational Rose 2000

PROFESSIONAL EXPERIENCE

April 2013
to Present

Quasi Governement

Ottawa, Ontario

HADOOP DEVELOPER

Data is being produced at a rate too large for relational databases. Lead developer in the transformation of data from relational DB to Hadoop.

Responsibilities:

- Developing parser and loader map reduce application to retrieve data from HDFS and store to HBase and Hive.
- Importing the data from the MySQL into the HDFS using Sqoop.
- Importing the unstructured data into the HDFS using Flume.
- Used Oozie to orchestrate the map reduce jobs that extract the data on a timely manner.
- Written Map Reduce java programs to analyze the log data for large-scale data sets.
- Involved in using HBase Java API on Java application.
- Automated all the jobs for extracting the data from different Data Sources like MySQL to pushing the result set data to Hadoop Distributed File System.
- Customize parser loader application of Data migration to HBase.
- Developed Pig Latin scripts to extract the data from the output files to load into HDFS.
- Developed custom UDFS and implemented Pig scripts.
- Implemented MapReduce jobs using Java API and PIG Latin as well HIVEQL
- Participated in the setup and deployment of Hadoop cluster
- Hands on design and development of an application using Hive (UDF).
- Responsible for writing Hive Queries for analyzing data in Hive warehouse using Hive Query Language (HQL).
- Provide support data analysts in running Pig and Hive queries.
- Involved in HiveQL.
- Involved in Pig Latin.
- Importing and exporting Data from MySQL/Oracle to HiveQL Using SQOOP.

- Importing and exporting Data from MySQL/Oracle to HDFS.
- Configured HA cluster for both Manual failover and Automatic failover.
- Designed and built many applications to deal with vast amounts of data flowing through multiple Hadoop clusters, using Pig Latin and Java-based map-reduce.
- Specifying the cluster size, allocating Resource pool, Distribution of Hadoop by writing the specification texts in JSON File format.
- Creates a SOLR schema from the Indexer settings
- Implemented SOLR index cron jobs.
- Experience in writing SOLR queries for various search documents
- Responsible for defining the data flow within Hadoop eco system and direct the team in implement them.
- Exported the result set from Hive to MySQL using Shell scripts.
- Developed HIVE queries for the analysts.

Environment: Apache Hadoop, Hive, Hue Tool, Zookeeper, Map Reduce, Sqoop, crunch API, Pig 0.10 and 0.11, HCatalog, Unix, Java, JSP, Eclipse, Maven, SQL, HTML, XML, Oracle, SQL Server, MYSQL

November 2008
to March 2013

Major Consulting Firm

Toronto, Ontario

BIG DATA DEVELOPER (January 2012 to March 2013)

Responsibilities:

- Handled importing of data from various data sources, performed transformations using Hive, PIG, and loaded data into HDFS.
- Experience in Importing and exporting data into HDFS and Hive using Sqoop.
- Load and transform large sets of structured, semi structured and unstructured data.
- Responsible for managing data coming from different sources.
- Gained good experience with NOSQL database.
- Involved in creating Hive tables, loading with data and writing hive queries, which will run internally in map, reduce way.
- Involved in creating tables, partitioning, bucketing of table.
- Good understanding and related experience with Hadoop stack-internals, Hive, Pig and Map/Reduce.

Environment: Core Java, MS Excel 2007, Oracle, Apache Hadoop, Pig, Hive, Map-reduce, Sqoop, JAVA/J2EE, WINDOWS.

Software Company

J2EE/JAVA DEVELOPER (February 2011 to December 2011)

Responsibilities:

- Involved in the designing of the project using UML.
- Followed J2EE Specifications in the project.
- Designed the user interface pages in JSP.
- Used XML and XSL for mapping the fields in database.
- Used JavaScript for client side validations.
- Created stored procedures and triggers that are required for project.
- Created functions and views in Oracle.
- Enhanced the performance of the whole application using the stored procedures and prepared statements.
- Responsible for updating database tables and designing SQL queries using PL/SQL.
- Created bean classes for communicating with database.
- Involved in documentation of the module and project.
- Prepared test cases and test scenarios as per business requirements.
- Involved in bug fixing.

- Prepared coded applications for unit testing using JUnit.