PRADEEP BHADANI

linkedin.com/in/pradeepbhadani
Citizenship: Indian
UK Work Status : Tier2 General VIsa

Phone: +44-7596703045
Mailto: pradeep.bhadani@gmail.com

"*Transforming large, unstructured data into useful information***"**

Profile
Big Data Engineer with 4.5years+ Hands-on experience in various Big Data Technologies.

Experience Summary
Big Data Partnership - As Big Data Engineer from **July,2014**
Cognilytics,Inc - As High Performance Computing Engineer from **January,2013 to June,2014**.
TCS Limited – As a Hadoop Developer From **July,2011 to December,2012**.

Certifications
- **edX Verified Certificate for Introduction to Big Data with Apache Spark**
- **IBM Big Data Spark Foundation Certificate**
- **Spark Fundamentals II - IBM Big Data University**
- **MapR Certified Hadoop Professional: Administrator (MCHP: A)**
- **MapR Certified Hadoop Professional: Developer (MCHP: D)**
- **Hortonworks Certified Apache Hadoop 2.0 Developer - Java**
- **Hortonworks Certified Apache Hadoop Administrator 2.x**
- **Cloudera Certified Apache Hadoop Developer (CCDH).**
- **Apache Drill Essentials & Architecture by MapR Academy**
- **Hadoop and Amazon Web Service by IBM Big Data University**
- **Hadoop Fundamentals I by IBM Big Data University**

Technical Skills

Big Data Technologies	Hadoop, HDFS, MapReduce, Spark ,PIG , Hive, Sqoop ,Kafka, Flume, HBase, Storm, HUE , ElasticSearch, Logstash, Kibana, Apache Tika, Snowtide, Ambari, Ranger, Knox
Tools / DevOps / Cloud	Eclipse, Chef , Ansible,Terraform, Vagrant , Serverspec, Behave, Cucumber , AWS (Amazon Web Services) , Google CLoud
Operating System	Linux , Windows
Languages / RDBMS	JAVA, PIG Latin, Shell Script, Python, MySQL

Professional Summary
- Worked in complete life-cycle of Software development.
- Extensive experience on building **50+** multi-node Hadoop Cluster (Hortonworks, MapR, Cloudera distribution) for various clients and internal use.
- Experience in building and managing AWS infrastructure using Terraform and Vagrant.
- Experience in Automating and deploying tools like Hortonworks Hadoop ,MapR Hadoop , Elasticsearch, Logstash, Kibana, Kafka using Chef , Ansible.
- Extensive Hands-on and Consulting experience in Hadoop EcoSystem like MapReduce, PIG , Hive, Hbase, Sqoop, Kafka etc..
- Exposure to different flavours of Hadoop - Cloudera, Hortonworks, MapR, IBM BigInsight, Apache.

- Trained more than **150+** professional on Hadoop - Cloudera, Hortonworks, MapR distributions.
- Exposure to other processing systems like Storm , Spark.
- Experience in Infrastructure and System Behaviour Testing using tech like serverspec , behave ,c cucumber.
- Text Extraction from PDF/Doc files using Apache tika and Snowtide.
- Experience in understanding customers multiple data sets which include Behavioral data, Sentiment analysis on social media data , Network data.
- Experience in Data transfer to & from different relational databases, legacy system like Mainframe to Hadoop Ecosystem.
- Develop MapReduce jobs, Hive UDFs, Hive UDAFs, PIG UDFs to perform Data Transformation and analysis.
- Writing ad-hoc analytic queries using PIG scripting , Hive HQL.
- Knowledge on cloud infrastructure like Amazon EC2.
- Good programming knowledge in Object Oriented language like Core Java.

Trainings
- Spark Developer Training.
- MapR training for Hadoop Developer and Admin
- Hortonworks Training for Developer , Operations.
- Intel Apache Hadoop (IDH) Training organized by Intel Corp.
- Amazon Web Services Training Programme organized by Amazon at TCS.
- Hadoop Internal Training Programme at TCS.

Responsibilities Undertaken in Big Data Partnership
Summary :
My role involved :
- Helping Architect team to choose big data technology , design solutions and develop POCs
- Automating Hadoop and other tools deployment using Chef/Ansible
- Building AWS infrastructure using Terraform and Vagrant.
- Building BDD and Infrastructure test for Hadoop systems
- Building and managing multi-node Hadoop Cluster , ELK cluster for different client.
- Developing application on Hadoop , ELK Stack
- Working closely with Client to gather requirement and understand changes in requirement.
- Helping client to understand Big data and need for different tool-set.

Responsibilities Undertaken in Cognilytics,Inc
Summary :
My role involved :
- Full-time Hadoop Consultant at VISA,Inc.
- Writing MR Jobs, Hive UDFs to accomplish different set of tasks.
- Providing Technical assistant to Business Users at VISA,Inc.
- Migrating from SAS to Hadoop.
- Enhancing the existing system using Automations at VISA,Inc.
- Working closely with Client.
- Writing Scripts to Manage Hadoop Clusters.

- Managing Hadoop clusters(100+ node) : install, monitor, maintain

Summary :

As a Team member & Developer my role involved :

- Doing R&D on Hadoop Framework , MapReduce and other related technologies.
- Developing Algorithm and Code for modules.
- Working with Market Analysis Team.
- Working as a member of Learning & Development Team in the project.

Project Undertaken :

Build multi-environment 12-node Hortonworks Secure Hadoop Production Cluster for Retail sector Client

Project is to build secure Hadoop Cluster for 4 different environment. Each environment consist of 12-node. Cluster was Kerberized , secured by Knox and Ranger and HA enabled.

- Write Scripts to build Hadoop Cluster (HA enabled).
- Build Ambari Blueprint for automation.
- Write BDD test for testing Hadoop Cluster.
- Kerberize Hadoop Cluster for Security.
- Enable Knox and Ranger for secure data access.
- Train client team to Admin Hadoop Cluster.

Upgrade Production Hadoop Cluster for two different clients

Project is to upgrade existing Hortonworks Hadoop cluster

- Plan upgrade to minimize downtime.
- Upgrade Hortonworks stack from HDP2.1 to HDP2.2.
- Install new components from new stack.
- Test and handover cluster to client.

Automate Hadoop Deployment

Deploying multi-node Hadoop is very complex and time consuming task.To simplify the deployment, organization needs the automated way to deploy clusters (like one-click installation).

- Develop Chef cookbook and Ansible Playbook for Big Data Technologies.
- Develop custom Ansible plugin as required
- Develop Infrastructure test using serverspec to smoothen the process of pre and post installation of Hadoop clusters.
- Develop BDD test using behave and cucumber to ensure Hadoop system works as expected. .

Data Lake for Europe's leading payment processor

Project is to migrate business process from legacy system to Hadoop.

- Deploy development and production Hadoop Cluster.
- Migrate data from legacy system to Hadoop cluster.
- Develop data pipelines to filter, convert and mask source data.

Monitoring System based on Logs for e-commerce player in UK

Project is to collect logs from different sources of an e-commerce business and draw meaningful insight. Used Elasticsearch to index data, Logstash to ingest data and Kibana to display graphs.

- Deploy ELK stack in production environment

- Develop Logstash configurations and filter for processing logs.
- Develop Kibana Dashboard as per requirement.

Develop Solution for Insurance sector client
Project was to migrate the Analytics system from Traditional to Big Data.
- Helped Client to identify the technologies to be used for the use-cases.
- Developed end-to-end Solution for the identified Use-cases and build Hadoop Cluster to deploy designed Solution.

Data Lake Solution for Digital Metadata Provider
Project is to migrate complete business logic and infrastructure for enriching Media metadata from Traditional system to Hadoop Ecosystem.
- Deploy MapR Hadoop in client production environment.
- Write MR jobs to transfer data from traditional system to HDFS.
- Clean source data and Map to Hive Tables.
- Transfer processed data to Elasticsearch for serving Dashboard.

Build 10-node Hortonworks Hadoop Production Cluster for Health sector Client
Project is to build 10-node Production Hadoop cluster for client. Cluster was Namenode HA enabled.
- Write Scripts to build Hadoop Cluster (Namenode HA enabled).
- Train client team to Admin Hadoop Cluster.

Build 5-node Hortonworks Hadoop Cluster for Public Sector Client
Build Hadoop Cluster on Scratch Machines.
- Install CentOS Operating System and configure client network on Cluster machines.
- Build Hadoop Cluster and transfer data to HDFS for further analytics.

Customer Analytics Model for US based Retailer
Customer Analytics Model aggregates the data coming from different sources (internal or external) like detail from POS , personal details etc.. Use this data to get Customer Insight view and serve them in a customized way.This model is used for other business purposes.Worked upon multiple modules :
- Develop Sqoop Jobs , Pig Scripts , MR jobs to transfer data from different storage systems to HDFS.
- Develop Hive UDFs and UDAFs to implement business logics.
- Develop MapReduce programs to transfer data from Hadoop Cluster to Cassandra Cluster.

Web Log Analysis
Project is to find loyal customers across the world and other customer pattern for Web Logs.Worked upon multiple modules :
- Develop Pig Script to analyse Unstructured Web access Logs.
- Develop PIG UDFs to perform data transformation to simplify the scripts.

Network Device Log Analyser :
Log Analyser generates Reports of Denied/Approved requests , Severity Level incidents etc. using Firewall and Server Logs.Worked upon following modules :
- Requirement Gathering
- Data Extraction using MapReduce programs , PIG Scripts, PIG UDFs , Regex.
- Develop Hive UDAFs for aggregations.

- Hadoop Cluster Setup (Cloudera Distribution).

Market Analysis on Big Data Analytics.

Done research about :
- Sectors having scope of Big Data Analytics.
- Technology Trend in the market.
- Vendors providing Products, Services and Support.
- Probable application of Big Data Analytics in various Sectors.

Visual Hbase : Visual Hbase is a web based user interface for Hbase.

The Web Application invokes the servlets to interact with Hbase using Java.

This application provides features for creating , reading and deleting a table. It also has feature for bulk insertion of data from CSV file without interacting with Hbase APIs.

Education Summary

Qualification Category	College/University	Subject
Bachelor Of Engineering	Institute of Engineering & Technology, DAVV	Computer Engineering