

<Name of Master Trainer> (Big Data, Data Science & Cyber Security)

QP/NOS Training Skills

- **Associate Analytics (All NOSs)**
- **Security Analyst (All NOSs)**
- **Junior Data Associate (All NOSs)**
- **Software Testing (All NOSs)**







Background

XXXX is a heading Big Data Infrastructure, having more than 14 years of professional experience Information Technology, expertise in mapping business problems to technical solutions and delivering customer value. Specialist in clusters setup, performance tuning and monitoring. Strong proponent of open source tools for solving tough business problems in Big Data. .

Education:

- Master of Science (Information Technology)
- Bachelor of Business Administration (BBA)

Certifications

-  : Cloudera Certified Administrator for Apache Hadoop (CCAH) CDH Version: 5
-  : AWS Certified Solutions Architect – Associate
-  : Data Science and Machine Learning Essentials
-  : IBM Certified Analyst – i2 Analysts Notebook V8.9.
-  : Cisco Certified Network Associate (CCNA R&S)
-  : Red Hat Certified Engineer (RHCE 5.0)

Expertise

Administration skills :

- Big Data - HDFS, MapReduce, Hive, Pig, Oozie, Nagios etc.
- RDBMS & NoSQL – MySQL, Hbase, MongoDB, Cassandra
- OS and Cloud – Linux, Windows, AWS, Good understanding over OpenStack
- Ethical Hacking (Trained)

Industry Experience:

- BFSI
- Telecom
- Banking
- Retail
- Healthcare
- Smart Cities

Selected Relevant Experience

Government Intelligence Agency found difficulty in running queries to identify closed group of user in SQL Server 2008, queries were running for 3 days and getting hanged on the Call data records Data. Client having 25TB of data consolidated from 16 different Mobile Operators across Andhra Pradesh Mobile Spectrum. Operators will give the data on weekly/fortnightly/monthly basis via external disk. Volume of each operator would be around 50 to 100GB per month.

- The challenge is to process multiple mobile operators and to identify the closed group of users (custom requirement).
- Setup a Hadoop cluster for cleansing and processing data with 24 nodes using IBM BladeCenter H.
- Installed and configured Hadoop 2.0.0 along with Echo system tools (Hive, Pig and Hbase) and performed data transformation into HDFS and then into Hive.
- Transformed all operators' data into a standard format in Hive for data cleansing.
- Loaded standardized data into HBase and ran MapReduce programs (Dev team developed a Graph Algorithm in order to find closed user groups by development team).
- Built production ready Hadoop cluster with 10 nodes on Amazon EC2 for Healthcare client.
- Built two clusters for PoCs for development team with 5 nodes each including one cluster on Amazon EC2 with NoSQL Databases HBase, MongoDB and Cassandra for development team.
- Implemented Nagios and Ganglia to monitor MR Job performance for the development team.
- Created multiple Virtual Machines for Pseudo and Multi node clusters on Amazon EC2/S3 with Hadoop Ecosystem tools for testing teams.
- Training BigData team on Hadoop Ecosystem and NoSQL Databases.
- Created various use cases using massive public data sets. Ran various performance tests for verifying the efficacy of MapReduce, Pig and Hive in various modes – standalone, pseudo distributed, cluster and Cloud. The exercise is part of a framework that is being developed to evaluate/assess various eco-system projects.
- Created various environments for aiding solution building like Development, test and production.
- Researched, collated, edited content related to Hadoop administration and delivered more than 100 hours of peer training.
- Very good understanding of Biometric devices, especially in Face Recognition.
- Major contributor for NASSCOMs for Skill development initiatives with Andhra Pradesh and Telangana State Governments for Jr. Data Associate, Associate Analyst and Security Analyst job roles.