

Bhanu C

Mobile: 770 728 8549

Hadoop Admin

E-mail ID: Krishna@tekniti.com

PROFESSIONAL SUMMARY:

- Total 5 years of IT experience and Extensive experience in the administration, modification, installation and maintenance of Hadoop on Linux RHEL operating system and Tableau.
- Expertise in designing Big Data Systems with Cloud Service Models - IaaS, PaaS, SaaS and FaaS (Serverless)
- Hands on experience in installation, configuration, supporting and managing Hadoop Clusters.
- In-depth knowledge of Hadoop Eco system - HDFS, Yarn, MapReduce, Hive, Hue, Sqoop, Flume, Kafka, Spark, Oozie, NiFi and Cassandra.
- Experience on Ambari (Hortonworks) for management of Hadoop Ecosystem.
- Expertise on setting up Hadoop security, data encryption and authorization using Kerberos, TLS/SSL and Apache Sentry respectively.
- Extensive hands on administration with Hortonworks.
- Practical knowledge on functionalities of every Hadoop daemon, interaction between them, resource utilizations and dynamic tuning to make cluster available and efficient.
- Designed and provisioned Virtual Network at AWS using VPC, Subnets, Network ACLs, Internet Gateway, Route Tables, NAT Gateways
- Strong knowledge on Hadoop HDFS architecture and Map-Reduce framework.
- Experienced in developing MapReduce programs using Apache Hadoop for working with Big Data.
- Good understanding of XML methodologies (XML, XSL, XSD) including Web Services and SOAP.
- Experience in administering the Linux systems to deploy Hadoop cluster and monitoring the cluster using Nagios and Ganglia.
- Experience in performing backup and disaster recovery of namenode metadata and important sensitive data residing on cluster.
- Architected and implemented automated server provisioning using puppet.
- Experience in performing minor and major upgrades.
- Experience in performing commissioning and decommissioning of data nodes on Hadoop cluster.
- Strong knowledge in configuring Name Node High Availability and Name Node Federation.
- Familiar with writing Oozie workflows and Job Controllers for job automation - shell, hive, Sqoop automation.
- Good working Knowledge in OOA and OOD using UML and designing use cases.
- Good understanding of Scrum methodologies, Test Driven Development and continuous integration.
- Familiar with importing and exporting data using Sqoop from RDBMS MySQL, Oracle, Teradata and using fast loaders and connectors Experience.
- Expertise in database performance tuning & data modeling.
- Experience in publishing Dashboards and presenting dashboards on web and desktop platforms to Tableau Server. Working Experience with system engineering team to plan and deploy Hadoop hardware and software environments.
- Storage/Installation, LVM, Linux Kickstart, Solaris Volume Manager, Sun RAID Manage.
- Expertise in Virtualizations System Administration of VMware EESX/EESXi, VMware Server, VMware Lab Manager, Vcloud, Amazon EC2 & S3 web services.
- Obtained experience on Commissioning, Decommissioning, Balancing, and Managing Nodes and tuning server for optimal performance of the ecosystem.
- Closely worked with Developers and Analysts to address project requirements. Ability to effectively manage time and prioritize multiple projects.

TECHNICAL SKILLS:

Hadoop /Big Data Technologies:HDFS, MapReduce, Hive, Pig, Sqoop, Flume, Oozie, Storm, Zookeeper, Kafka, Impala, HCatalog, Apache Spark, Spark Streaming, Spark SQL, HBase, NiFi and Cassandra, AWS (EMR, EC2), Horton Works, Cloudera

Languages:Java, SQL

Protocols:TCP/IP, HTTP, LAN, WAN

Network Services:SSH, DNS/BIND, NFS, NIS, Samba, DHCP, Telnet, FTP, IPtables, MS AD/LDS/ADC and OpenLdap.
Other Tools Tableau, SAS

Mails Servers and Clients:Microsoft Exchange, Lotus Domino, Send mail, Postfix.

Databases:Oracle 9g/10g & MySQL 4.x/5.x, HBase, NoSQL, Postgres

Platforms:Red Hat Linux, Centos, Solaris, and Windows

Software Methodologies:Agile Methodology -SCRUM, Hybrid

PROFESSIONAL EXPERIENCE:

Client: GAIG, Cincinnati, OH.

Role: Hadoop AdminMar 2019 - Present

Responsibilities:

- Worked on analyzing Hortonworks Hadoop cluster and different big data analytic tools including Pig, HBase Database and Sqoop.
- Responsible for building scalable distributed data solutions using Hadoop.
- Experienced on setting up Horton works cluster and installing all the ecosystem components through Ambari and manually from command line.
- Performed a Major upgrade in production environment from HDP 1.3 to HDP 2.2 and followed standard Back up policies to make sure the high availability of cluster.
- Monitored multiple Hadoop clusters environments using Ganglia and Nagios. Monitored workload, job performance and capacity planning using Ambari.
- Responsible for Cluster maintenance, Monitoring, commissioning and decommissioning Data nodes, Troubleshooting, Cluster Planning, Manage and review data backups, Manage & review log files.
- Experienced on setting up Horton works cluster and installing all the ecosystem components through Ambari and through command line interface.
- Developed scripts for tracking the changes in file permissions of the files and directories through audit logs in HDFS.
- Configured memory and v-cores for the dynamic resource pools within the fair and capacity scheduler.
- Implemented test scripts to support test driven development and continuous integration.
- Worked intact for integrating the LDAP server and active directory with the Ambari through command line interface.
- Coordinated with Hortonworks support team through support portal to sort out the critical issues during upgrades.
- Worked on smoke tests of each service and client upon installation and configuration. Involved in loading data from UNIX file system to HDFS.
- Load and transform large sets of structured, semi structured and unstructured data. Cluster coordination services through Zookeeper.
- Experience in managing and reviewing Hadoop log files. Installed Oozie workflow engine to run multiple Map Reduce, Hive and pig jobs.
- Installed Knox gateway on a separate node for providing REST API services and to give HDFS access to various http applications.
- Balancing HDFS manually to decrease network utilization and increase job performance.
- Commission and decommission the Data nodes from cluster in case of problems.
- Set up automated processes to archive/clean the unwanted data on the cluster, in particular on HDFS and Local file system.
- Set up and manage High Availability of Name node federation using Quorum Manager to avoid single point of failures in large clusters.

Environment: Hortonworks 2.2.1, HDFS, Hive, Pig, Sqoop, HBase, Micro strategy, Shell Scripting, Ubuntu, RedHat Linux.

Client: Equifax, Atlanta,GA.

Role: Hadoop Admin July 2018 - Mar 2019

Responsibilities:

- Responsible to drive and fix any production "severity one" from technical standpoint.
- Manage over ~2500 Hadoop ETL jobs in production. Manage Production cluster comprises of 220 nodes.
- Involved in deploying a Hadoop cluster using Hortonworks Ambari HDP 2.2 integrated with Sitescope for monitoring and Alerting.
- Launching and Setup of Hadoop Cluster on AWS as well as physical servers, which includes configuring different components of Hadoop.
- Created a local YUM repository for installing and updating packages. Configured and deployed hive metastore using MySQL and thrift server.
- Responsible for building system that ingests terabytes of data per day into Hadoop from a variety of data sources providing high storage efficiency and optimized layout for analytics.
- Developed data pipelines that ingests data from multiple data sources and process them.
- Expertise in Using Sqoop to connect to the ORACLE, MySQL, SQL Server, TERADATA and move the pivoted data to Hive tables or Hbase tables.
- Configured Kerberos for authentication, Knox for perimeter security and Ranger for granular access in the cluster.
- Configured and installed several Hadoop clusters in both physical machines as well as the AWS cloud for POCs.
- Developed Simple to complex MapReduce Jobs using Hive and Pig. Involved in creating Hive tables and loading and analyzing data using hive queries
- Extensively used Sqoop to move the data from relational databases to HDFS. Used Flume to move the data from web logs onto HDFS.
- Used Pig to apply transformations validations, cleaning and deduplication of data from raw data sources.
- Used Storm service extensively to connect to Active MQ and KAFKA to push data to HBASE and HIVE tables.
- Used NIFI to pull the data from different source and to push the data to HBASE and HIVE
- Worked on installing SPARK and performance tuning. Upgraded the Hadoop cluster from HDP 2.2 to HDP 2.4
- Integrated schedulers Tidal and Control-M with the Hadoop clusters to schedule the jobs and dependencies on the cluster.
- Worked closely with the Continuous Integration team to setup tools like Github, Jenkins and Nexus for scheduling automatic deployments of new or existing code.
- Actively monitored the Hadoop Cluster of 220 Nodes with Hortonworks distribution with HDP 2.4.
- Performed various configurations, which includes, networking and IPTable, resolving hostnames, user accounts and file permissions, http, ftp, SSH keyless login.
- Have worked on installing Hadoop services on cloud integrated with AWS. Integrated BI tool Tableau to run visualizations over the data.
- Worked closely with the Continuous Integration team to setup tools like Github, Jenkins and Nexus for scheduling automatic deployments of new or existing code.
- Provided 24 x 7 on call support as part of a scheduled rotation with other team members.

Environment: Hadoop HDFS, Map Reduce, Hive, Pig, Oozie, Sqoop, Ambari, Nifi, Storm, Aws S3, Ec2, Iam, Zookeeper, Splunk, Kafka.

Client: Mastercard, Hyderabad, India.

Role: Hadoop Admin Aug 2015–Dec 2016

Responsibilities:

- Currently working as admin on Cloudera (CDH 5.5.1) distribution for 4 clusters ranges from POC to PROD. Installed Oozie workflow engine to run multiple Hive and pig jobs.

- Responsible for Cluster maintenance, Monitoring, commissioning and decommissioning Data nodes, Troubleshooting, Manage and review data backups, Manage & review log files.
- Adding/installation of new components and removal of them through Cloudera Manager. Involved in extracting the data from various sources into Hadoop HDFS for processing.
- Collaborating with application teams to install operating system and Hadoop updates, patches, version upgrades.
- Experience with cloud AWS/EMR, Cloudera Manager (also direct- Hadoop -EC2(non EMR))
- Monitored workload, job performance and capacity planning using Cloudera Manager. Rack Aware Configuration and AWS working nature
- Involved in Analyzing system failures, identifying root causes and recommended course of actions. Imported logs from web servers with Flume to ingest the data into HDFS.
- Interacting with Cloudera support and log the issues in Cloudera portal and fixing them as per the recommendations. Fine tuning hive jobs for optimized performance.
- Working experience in supporting and deploying in an AWS environment. Using Flume and Spool directory loading the data from local system to Hdfs
- Retrieved data from HDFS into relational databases with Sqoop. Parsed cleansed and mined useful and meaningful data in HDFS using Map-Reduce for further analysis
- Implemented custom interceptors for flume to filter data and defined channel selectors to multiplex the data into different sinks.
- Worked on analyzing Hadoop cluster and different big data analytic tools including Pig, Hbase database and Sqoop. Partitioned and queried the data in Hive for further analysis by the BI team.
- Configuring, Managing permissions for the users in hue. Responsible for building scalable distributed data solutions using Hadoop.
- Commissioned and Decommissioned nodes on CDH5 Hadoop cluster on Red hat LINUX.
- Creating and managing the Cron jobs. Extending the functionality of Hive and Pig with custom UDF's and UDAF's. Implemented test scripts to support test driven development and continuous integration.
- Worked on tuning the performance Pig queries. Creating and truncating Hbase tables in hue and taking backup of submitter ID(s).
- Worked with application teams to install operating system, Hadoop updates, patches, version upgrades as required. Involved in loading data from LINUX file system to HDFS.
- Experience in configuring the Storm in loading the data from MYSQL to HBASE using jms
- Responsible to manage data coming from different sources. Experience in managing and reviewing Hadoop log files. Involved in loading data from UNIX file system to HDFS.
- Exported the analyzed data to the relational databases using Sqoop for visualization and to generate reports for the BI team.
- Analyzed large amounts of data sets to determine optimal way to aggregate and report on it.
- Supported in setting up QA environment and updating configurations for implementing scripts with Pig and Sqoop.

Environment: Hdfs, Map reduce, Hive 1.1.0, Hue 3.9.0, Pig, Flume, Oozie, Sqoop, Cdh5, Apache Hadoop 2.6, Spark, Solr, Storm, Cloudera Manager, Red hat, MySQL and Oracle.

Client: Fortune Brands Home and Security, Hyderabad, India.

Role: Hadoop Developer

May 2014 - Aug 2015

Responsibilities:

- Involved in architecture design, development and implementation of Hadoop deployment, backup and recovery systems
- Responsible for Installation and configuration of on the Hortonworks Hadoop cluster using Ambari
- Developed Sqoop scripts to import data from Oracle database and handled incremental loading on the point of sale tables
- Experience in writing Pig scripts to transform raw data from several data sources into forming baseline data
- Created Hive external tables, views and scripts for transformations such as filtering, aggregation and partitioning tables

- Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting
- Implementing different performance optimization techniques such as using distributed cache for small datasets, partitioning and bucketing in Hive, doing map side joins etc.
- Developed User Defined Functions (UDF) in Java if required for Hive queries
- Developed bash scripts to automate the above process of Extraction, Transformation and Loading
- Managing Hadoop cluster using Ambari and created roles and user groups in Ambari for permitted access to Ambari functions
- Working knowledge on ZooKeeper

Client: JDS Uniphase Corporation, Hyderabad, India.

Role: DatabaseDeveloper(Unpaid Internship)April 2012 – Sep 2013

Responsibilities:

- Developed SQL scripts to Insert/Update and Delete data
- Written complex SQL statements using joins, scalar and table user defined functions, views
- Worked on stored procedures and database triggers and worked
- Generated database SQL scripts and deployed databases including installation and configuration
- Experience in creating Indexes for faster performance and views for controlling user access to data
- Develop Logical/Physical Model for both Transactional and Data Warehouse systems
- Involved in Analysis, design, development & implementation of FACT tables, Dimension tables using Star schema

Environment: SQL Server 2005, DDL, DML, Store Procedures, Views, Waterfall methodology.

BACHELOR'S IN ELECTRONICS AND COMMUNICATIONS- Gandhi Institute of Technology - India - 2010-2014

MASTER OF SCIENCE IN COMPUTER SCIENCE- University of New Haven-West Haven, CT. May 2018

Bhanu Chand

- Broomfield, CO, USA

Contact Information

- conjava8@gmail.com (Preferred)
- 7707288549 (Preferred)

Work History

- **Hadoop Admin | Gai**
Jan 01, 2019 - No End Date

Skills

- **configuration** | 7yrs | 2019
- **scripting** | 6yrs | 2019
- **installation** | 6yrs | 2019
- **apache pig** | 5yrs | 2019
- **cluster** | 5yrs | 2019
- **apache hadoop** | 5yrs | 2019
- **apache hive** | 5yrs | 2019
- **apache sqoop** | 5yrs | 2019
- **continuous integration** | 4yrs | 2019
- **database** | 4yrs | 2019
- **mapreduce** | 4yrs | 2019
- **oracle** | 4yrs | 2019
- **apache hbase** | 4yrs | 2019
- **apache oozie** | 4yrs | 2019
- **hdfs** | 4yrs | 2019
- **monitoring** | 4yrs | 2019
- **apache flume** | 3yrs | 2019
- **apache spark** | 3yrs | 2019
- **apache zookeeper** | 3yrs | 2019
- **hortonworks** | 3yrs | 2019
- **apache**
- **hadoop**

- hbase
- hive

Work Preferences

- Likely to Switch: Most Likely
- Willing to Relocate: Yes
- Work Authorization:
 - US
- Work Documents:
 - Employment Auth Document
- Security Clearance: No
- Third Party: Yes
- Employment Type:
 - Contract - W2
 - Contract to Hire - Independent
 - Contract to Hire - W2
 - Contract - Independent

Profile Sources

- Dice:
<https://www.dice.com/employer/talent/profile/d74df520b5c13ec9a139b14ca1574794c7cc64a4>