

# BHARAT KATRAGADDA

---

Email: [katragaddabharat6@gmail.com](mailto:katragaddabharat6@gmail.com) Cell: 5135939206

## Summary:

- Over 6+ years of total Information Technology experience with expertise in Administration and Operations experience, in Big Data and Cloud Computing Technologies
- Expertise in setting up fully distributed multi node (Bigdata)Hadoop clusters, with Apache, Cloudera Hadoop.
- Expertise in AWS services such as EC2, Simple Storage Service (S3), Autoscaling, EBS, Glacier, VPC, ELB, RDS, IAM, Cloud Watch, and Redshift.
- Expertise in MIT kerberos and High Availability as well as Integration of Hadoop clusters.
- Experience in upgrading Hadoop clusters.
- Strong knowledge in installing, configuring and using ecosystem components like Hadoop MapReduce, Oozie, Hive, Sqoop, Pig, Flume, Zookeeper, Kafka, NameNode Recovery, HDFS High Availability Experience in Hadoop Shell commands, verifying managing and reviewing Hadoop Log files.
- Experienced with Devops tools like Chef, Puppet, Ansible, Jenkins, Jira, Docker and Splunk.
- Designed and Implemented CI & CD Pipelines achieving the end to end automation Supported server/VM provisioning activities, middleware installation and deployment activities via puppet.
- Written puppet manifests Provision several pre-prod environments.
- Written puppet modules to automate our build/deployment process and do an overall process improvement to any manual processes.
- Designed, Installed and Implemented / puppet. Good Knowledge in automation by using Puppet
- Implementing AWS architectures for web applications
- Experience in EC2, S3, ELB, IAM, Cloudwatch, VPC in AWS
- Experience in understanding the security requirements for Hadoop and integrate with Kerberos authentication and authorization infrastructure.
- Extensive experience on performing administration, configuration management, monitoring, debugging, and performance tuning in Hadoop Clusters.
- Performed AWS EC2 instance mirroring, WebLogic domain creations and several proprietary middleware Installations.
- Worked in agile projects delivering end to end continuous integration/continuous delivery pipeline by Integration of tools like Jenkins, puppy and AWS for VM provisioning.
- Evaluating performance of EC2 instances their CPU, memory usage and setting up EC2 Security Groups and VPC.
- Configured and Managed Jenkins in various Environments, Linux and Windows.
- Administered Version Control systems GIT, to create daily backups and checkpoint files.
- Created various branches in GIT, merged from development branch to release branch and created tags for releases.
- Experience creating, managing and performing container based deployments using Docker images Containing Middleware and Applications together.
- Enabling/Disabling of Passive and Active check for Hosts and Service in Nagios.
- Good knowledge in installing, configuring & maintaining Chef server and workstation
- Expertise in provisioning clusters and building manifests files in puppet for any services.

- Excellent knowledge in Import/Export structured, un-structured data from various data sources such as RDBMS, Event logs, Message queues into HDFS, using a variety of tools such as Sqoop, Flume etc.
- Expertise in converting non kerberized Hadoop cluster to Hadoop with kerberized cluster
- Administration and Operations experience with Big Data and Cloud Computing Technologies
- Handling in setting up fully distributed multi node Hadoop clusters, with Apache and AWS EC2instances
- Handling in AWS services such as EC2, Simple Storage Service(S3), Auto scaling, EBS, ELB, RDS, IAM, Cloud Watch
- Performing administration, configuration management, monitoring, debugging, and performance tuning in Hadoop Clusters.
- Authorized to work in United States for any employer

## Technical Skills:

- **Monitoring Tools:** Cloudera Manager, Ambari, Nagios, Ganglia
- **Scripting Languages:** Shell Scripting, Puppet, Scripting, Python, Bash, CSH.
- **Hadoop Distribution:** Horton works Distribution Platform 2.5, Cloudera Distribution.
- **Programming Languages:** C, Java, SQL, and PL/SQL.
- **Front End Technologies:** HTML, XHTML, XML.
- **Application Servers:** Apache Tomcat, Web Logic Server, Web sphere
- **Databases:** Oracle 11g, MySQL, MS SQL Server, IBM DB2.
- **NoSQL Databases:** HBase, Cassandra, MongoDB
- **Operating Systems:** Linux, UNIX, MAC, Windows NT / 98 /2000/ XP/ Vista, Windows 7, Windows 8.
- **Networks:** HTTP, HTTPS, FTP, UDP, TCP/TP, SNMP, SMTP.
- **Security:** Kerberos, Ranger, Knox, Falcon.

## Work- Experience:

### Hadoop Administrator

Avangrid –Orange,CT

Oct 2018 to Present

### Responsibilities:

- Worked on Distributed/Cloud Computing (MapReduce/ Hadoop(Bigdata), Hive, Pig, HBase, Sqoop, Flume, Spark, Zookeeper, etc.), Hortonworks (HDP 2.5.0).
- Deploying, managing, and configuring HDP using Apache Ambari 2.4.2.
- Installing and Working on Hadoop clusters for different teams, supported 100+ users to use Hadoop platform and resolve tickets and issues they run into and provide training to users to make Hadoop usability simple and updating them for best practices.
- Installed/Configured/Maintained Apache Hadoop(Bigdata) clusters for application development and Hadoop tools like Hive, Pig, HBase, Zookeeper and Sqoop.
- Effectively using NIFI expression language to enrich flow file attributes.
- Configuring YARN capacity scheduler with Apache Ambari.
- Configuring predefined alerts and automating cluster operations using Apache Ambari.
- Managing files on HDFS via CLI/Ambari files view.
- Responsible for efficient operations of multiple Cassandra clusters.
- Ensure the cluster is healthy and availablewith monitoring tool.

- Built high availability for major production cluster and designed automatic failover control using Zookeeper Failover Controller (ZKFC) and Quorum Journal nodes.
- Implemented Flume, Spark, Spark Stream framework for real time data processing.
- Involved in implementing security on Hortonworks Hadoop Cluster using with Kerberos by working along with operations team to move non-secured cluster to secured cluster.
- Responsible for upgrading Hortonworks Hadoop HDP2.2.0 and MapReduce 2.0 with YARN in Multi Clustered Node environment.
- Working with Talend to Loading data into (Bigdata)Hadoop Hive tables and Performing ELT aggregations in Hadoop Hive and also Extracting data from Hadoop Hive.
- Responsible for services and component failures and solving issues through analyzing and troubleshooting the Hadoop cluster.
- Manage and review Hadoop log files. Monitor the data streaming between web sources and HDFS.
- Managing Ambari administration, and setting up user alerts.
- Handled importing of data from various data sources, performed transformations using Hive, MapReduce, Spark and loaded data into HDFS.
- Solving Hive thrift issues and HBase problems after upgrading HDP 2.4.0.
- Involved in projects Extensively on Hive, Spark, Pig, Sqoop and Gemfire XD throughout the development Lifecycle until the projects went into Production.
- Managing the cluster resources by implementing capacity scheduler by creating queues.
- Integrated Kafka with Flume in sand box Environment using Kafka source and Kafka sink.
- Performed Puppet, Kibana, Elastic Search, Talend, Red Hat infrastructure for data ingestion, processing, and storage.
- Monitored multiple Hadoop clusters environments using Ganglia and Nagios. Monitored workload, job performance and capacity planning using Ambari.
- Performed many complex system analyses to improve ETL performance, identified high critical batch jobs to prioritize.
- Patch, Upgrade and keep the PostgreSQL DBs current. Develop and enhance scripts to automate and execute various DBA tasks.
- Implemented Spark solution to enable real time reports from Hadoop data. Was also actively involved in designing column families for various Hadoop Clusters.

**Environment:** HDP 2.5.0, Ambari 2.4.2, Oracle 11g/10g, MySQL, Sqoop, Teradata, Hive, Oozie, Spark, ZooKeeper, Talend, MapReduce, Apache NiFi, Pig, Kerberos, RedHat 7.

## **Hadoop Administrator**

**Intuit-San Jose, CA**

**Oct 2016 to Sep 2018**

### **Responsibilities:**

- Administration & Monitoring Hadoop.
- Worked on Hadoop Upgradation from 4.5 to 5.2
- Monitor Hadoop cluster job performance and capacity planning
- Removing from monitoring of particular security group nodes in nagios in case of retirement
- Responsible for managing and scheduling jobs on Hadoop Cluster
- Monitoring Hadoop Cluster through Cloudera Manager and Implementing alerts based on Error messages. Providing reports to management on Cluster Usage Metrics and Charge Back customers on their Usage.

- Splunk configuration that involves different web application and batch, create Saved search and summary search, summary indexes.
- Replacement of Retired Hadoop slave nodes through AWS console and Nagios Repositories
- Performed dynamic updates of Hadoop Yarn and MapReduce memory settings
- Nagios Xi Applications monitoring thru SNMP MIBs and OIDs.
- Worked with DBA team to migrate Hive and Oozie meta store Database from MySQL to RDS
- Worked with fair and capacity schedulers, creating new queues, adding users to queue, Increase mapper and reducers capacity and also administer view and submit MapReduce jobs
- Experience in Administration/Maintenance of source control management systems, such as GIT and GITHUB knowledge
- Hands on experience in installing and administrating CI tools like Jenkins
- Experience in integrating Shell scripts using Jenkins
- Installed and configured an automated tool Puppet that included the installation and configuration of the Puppet master, agent nodes and an admin control workstation.
- Working with Modules, Classes, Manifests in Puppet
- Experience in creating Docker images
- Used containerization technologies like Docker for building clusters for orchestrating containers deployment.
- Operations - Custom Shell scripts, VM and Environment management.
- Experienced in working with Amazon EC2, S3, Glaciers
- Create multiple groups and set permission polices for various groups in AWS
- Experienced in creating life cycle policies in AWS S3 for backups to Glaciers
- Experienced in maintaining, executing, and scheduling build scripts to automate DEV/PROD builds.
- Configured Elastic Load Balancers with EC2 Auto scaling groups.
- Created monitors, alarms and notifications for EC2 hosts using Cloudwatch.
- Launching Amazon EC2 Cloud Instances using Amazon Images (Linux/Ubuntu) and configuring launched instances with respect to specific applications
- Worked with IAM service creating new IAM users & groups, defining roles and policies and Identity providers
- Experienced in assigning MFA in AWS using IAM and s3 buckets
- Defined AWS Security Groups which acted as virtual firewalls that controlled the traffic allowed to reach one or more AWS EC2 instances.
- AmazonRoute53 to oversee DNS zones and furthermore give open DNS names to flexible load balancers IP.
- Using default and custom VPCs to create private cloud environments with public and private subnets
- Loaded data from Oracle, MS SQL Server, MySQL, Flat File database into HDFS, HIVE
- Fixed Namenode partition failed, fsimage not rotated, MR job failed with too many fetch failures and troubleshooting common Hadoop cluster issues
- Implemented manifest files in puppet for automated orchestration of Hadoop and Cassandra clusters
- Maintaining Github repositories for Configuration Management
- Configured distributed monitoring system Ganglia for Hadoop clusters
- Managing cluster coordination services through Zoo Keeper
- Configured and deployed Namenode High Availability Hadoop cluster with SSL and kerberoized
- Deal with the several services restart and killing the process with Pid to clear the alert
- Monitoring Log files of several services, clear files incase of Diskspace issues on share this nodes
- 24X7 production support for weekly schedule with Ops team

**Environment:**CentOS, CDH4, Hive, Sqoop, Flume, Hbase, MySQL, Cassandra, Oozie, Puppet, Pager Duty, Nagios, AWS (S3, EC2, IAM, Cloud Watch, RDS, ELB, Auto Scaling, EBS, VPC, EMR, Github.

## **Hadoop Administrator**

**Cisco- San Jose, CA**

**July 2015 to Sep 2016**

### **Responsibilities:**

- Responsible for building scalable distributed data solutions using Hadoop.
- Involved in Installing and configuring Hive, Pig, Sqoop, Flume and Oozie on the Hadoop cluster.
- Handled importing of data from various data sources, performed transformations using Hive, MapReduce, loaded data into HDFS and Extracted the data from MySQL into HDFS using Sqoop.
- Hands on experience in Hadoop administration and support activities for installations and configuring Apache Big Data Tools and Hadoop clusters using Cloudera Manager.
- Capable to handle Hadoop cluster installations in various environments such as Unix, Linux and Windows, able to implement and execute Pig Latin scripts in Grunt Shell.
- Experienced with file manipulation, advanced research to resolve various problems and correct integrity for critical Big Data issues with NoSQL Hadoop HDFS Database.
- Collected the logs data from web servers and integrated in to HDFS using Flume.
- Implemented NameNode backup using NFS. This was done for High availability.
- Developed PIG Latin scripts to extract the data from the web server output files to load into HDFS.
- Involved in the installation of CDH3 and up-gradation from CDH3 to CDH4.
- Created Hive External tables and loaded the data in to tables and query data using HQL.
- Worked with application teams to install operating system, Hadoop updates, patches, version upgrades as required.
- Involved in Installing the Oozie workflow engine in order to run multiple Hive and Pig jobs.
- Exported the analyzed data to the relational databases using Sqoop for visualization and to generate reports for the BI team.

**Environment:**Hadoop, MapReduce, HDFS, Hive, Pig, Java, SQL, Cloudera Manager, Sqoop, Flume, Oozie, CDH3, MongoDB, Cassandra, HBase, Java, Eclipse, Oracle and Unix/Linux.

## **Linux System Administrator**

**Advanced Integrated Research Solutions, Hyderabad**

**June 2012 to May 2014**

### **Responsibilities:**

- Monitored everyday systems and evaluated availability of all server resources and performed all activities for Linux servers.
- Installed and maintained all server hardware and software systems and administered server performance and ensured availability for same.
- Responsible for creating and managing users accounts, groups and security policies.
- Troubleshooting network issues and system maintenance, resolving software and hardware issues.
- Provided support of physical servers and virtual servers in a production environment.
- Created users account, adding /removing users, password reset, updating users profile, setting permissions on files and directories.
- Good Experience in setting up the Linux environments, Password less SSH, creating file systems, disabling firewalls, swapping, Selinux and installing Java.
- Creating new file system, managing & checking data consistency of the file system.

- Ability to diagnose network problems and understood TCP/IP networking and its security considerations.
- Monitored and Log Management on RHEL Centos, Ubuntu servers including processes, crash dumps and swap management, with password recovery.
- Adding, removing, or updating user account information and resetting passwords.
- Using Java JDBC to load data into MySQL.
- Maintaining the MySQL server and Authentication to required users for database access.
- Installing and updating packages using YUM.
- Patches installation and updating on server.
- Installation and configuration of Linux for new build environment.
- Did volume management using LVM and creating of physical and logical volume groups.
- Hands-on experience in Linux admin activities on RHEL, Cent OS & Ubuntu.
- Excellent in communicating with clients, customers, managers, and other teams in the enterprise at all levels.
- Effective problem-solving skills and outstanding interpersonal skills.
- Ability to work independently as well as within a team environment and driven to meet deadlines.
- Motivated to produce robust and high-performance output work.
- Ability to learn and use new technologies quickly.

**Environment:** Oracle Red Hat Linux6, 7; Linux, Centos, Ubuntu, VMware, LVM, TCP/IP, MYSQL

# BHARAT KATRAGADDA

- Orange, CT, USA

## Contact Information

- katragaddabharat6@gmail.com (Preferred)
- 5135939206 (Preferred)

## Work History

### Total Work Experience: 8 years

- **Hadoop Administrator | Capital One**  
Oct 01, 2017 - No End Date | Los Angeles CA United States
- **Hadoop Administrator | Intuit**  
Oct 01, 2016 - Sep 01, 2017 | San Jose CA United States
- **Hadoop Administrator | Advanced Integrated Research Solutions**  
Jul 01, 2013 - Sep 01, 2016
- **Linux System Administrator | Soft standard Solutions**  
Jun 01, 2011 - May 01, 2013

## Skills

- **mysql** | 6yrs | 2019
- **configuration** | 6yrs | 2019
- **security** | 5yrs | 2019
- **troubleshooting** | 5yrs | 2019
- **cluster** | 4yrs | 2019
- **mapreduce** | 4yrs | 2019
- **apache cassandra** | 4yrs | 2019
- **apache flume** | 4yrs | 2019
- **apache hadoop** | 4yrs | 2019

- **apache hbase** | 4yrs | 2019
- **apache hive** | 4yrs | 2019
- **apache oozie** | 4yrs | 2019
- **apache sqoop** | 4yrs | 2019
- **hdfs** | 4yrs | 2019
- **database** | 5yrs | 2018
- **linux** | 5yrs | 2018
- **installation** | 5yrs | 2018
- **software** | 5yrs | 2018
- **centos** | 4yrs | 2018
- **ubuntu** | 4yrs | 2018
- **virtual machines** | 4yrs | 2018
- **apache**
- **cassandra**
- **hadoop**
- **hbase**
- **hive**

## Work Preferences

- Likely to Switch: Most Likely
- Willing to Relocate: Yes
- Work Authorization:
  - US
- Work Documents:
  - Have H1
- Security Clearance: Yes
- Third Party: Yes
- Employment Type:
  - Contract - Corp-to-Corp
  - Contract - W2
  - Contract to Hire - Independent
  - Contract to Hire - W2
  - Contract - Independent
  - Contract to Hire - Corp-to-Corp

## Profile Sources

- Dice:
   
<https://www.dice.com/employer/talent/profile/7171b8cbc63fee8a6dfc46c44076fe29c9ff297b>