

# Manikya Vardhan

## Hadoop Admin

Boston MA - 01821 / (209) 813 1243 / Manikasturi5@gmail.com

Overall around 6+ years of professional Information Technology experience including 5+ years of Hadoop and Linux Administration activities such as installation, configuration and maintenance of systems/clusters. Hands on day-to-day operation of the environment, knowledge and deployment experience in Hadoop ecosystem.

Extensive knowledge and experience in Big Data with Map-Reduce, HDFS, Hive, Pig, Impala, Spark, Kafka, Zookeeper, Oozie, Sentry, HUE and Sqoop. Good knowledge of Hadoop Architecture and various components such as HDFS, Job Tracker, Task Tracker, Resource Manager, Name Node, Data Node, and Map Reduce(MRV1 and YARN) concepts.

Experience in performing various major and minor Hadoop upgraded on large environments and good exposure with Securing Hadoop clusters using Kerberos.

- Hands on experience in Hadoop Clusters using Horton works (HDP), Cloudera (CDH4, CDH5), oracle big data and Yarn distributions platforms.
- Responsible for implementation and ongoing administration of Hadoop infrastructure.
- Strong exposure to IT consulting, software project management, team leadership, design, development, implementation, maintenance/support and Integration of Enterprise Software Applications.
- Experience in Big Data processing using Apache Hadoop. Extensive experience in HDFS, MapReduce, Pig and Hive.
- Primary technical skills in Apache Hadoop, Map-Reduce, Pig, Hive, Hbase, ZooKeeper, Sqoop, Flume, OOZIE, Core Java, Java Script, J2EE, Oracle 11G/10G, HP PPM 9.2.
- Experience in Databases such as Oracle 11G/10G, DB2, Teradata, MySQL and Sybase.
- Having working knowledge on Windows and UNIX shell scripting.
- Aligning with the systems engineering team to propose and deploy new hardware and software environments required for Hadoop and to expand existing environments.
- Working with data delivery teams to setup new Hadoop users. This job includes setting up Linux users, setting up Kerberos principals and testing HDFS, Hive, Pig and MapReduce access for the new users.
- Cluster maintenance as well as creation and removal of nodes using tools like Ganglia, Nagios, Cloudera Manager Enterprise, Dell Open Manage and other tools.
- Performance tuning of Hadoop clusters and Hadoop MapReduce routines.
- Screen Hadoop cluster job performances and capacity planning.
- Monitor Hadoop cluster connectivity and security.
- Manage and review Hadoop log files.
- File system management and monitoring.
- HDFS support and maintenance.
- Diligently teaming with the infrastructure, network, database, application and business intelligence teams to guarantee high data quality and availability.
- Collaborating with application teams to install operating system and Hadoop updates, patches, version upgrades when required.
- Knowledge of Data Ware Housing concepts and Cognos 8 BI Suit and Business Objects.

### **Technical Skills:**

- **Big Data Ecosystems:** Hadoop, Map Reduce, HDFS, HBase, Zookeeper, Hive, Pig, Sqoop, Kafka, Cassandra, Oozie, Flume and Talend
- **Cloud:** AWS (EC2, S3, ELB, VPC, Auto Scaling), Azure
- **Programming Languages:** Java, C/C++, Assembly Language (8085/8086)
- **Scripting Languages:** JSP & Servlets, JavaScript, XML, HTML, Python and shell
- **Databases:** NoSQL, Oracle
- **UNIX Tools:** Apache, Yum, RPM
- **Tools:** Eclipse, JDeveloper, JProbe, CVS, Ant, MS Visual Studio
- **Platforms:** Windows (2000/XP), Linux, Solaris, AIX, HP/UX

- **Application Servers:** Apache Tomcat 5.x 6.0, Jboss 4.0
- **Testing Tools:** Net Beans, Eclipse, WSAD
- **Methodologies:** Agile

## **Professional Experience**

---

**Nov 2018 – CURRENT**

**Hadoop Administrator, Insulet Corporation  
Boston, MA**

### **Responsibilities:**

- Worked as Hadoop Admin and responsible for taking care of everything related to the clusters total of 100 nodes ranges from POC (Proof-of-Concept) to PROD clusters.
- Responsible for Cluster Maintenance, Monitoring, Managing, Commissioning and decommissioning Data nodes, Troubleshooting, and review data backups, Manage & review log files.
- Responsible for upgrading MapReduce 2.0 with YARN in Multi Clustered Node environment.
- Handled importing of data from various data sources, performed transformations using Hive, Map Reduce, Spark and loaded data into HDFS.
- Hadoop security setup using MIT Kerberos, AD integration (LDAP) Sentry authorization.
- Installed and configured Kafka Cluster.
- Configured various property files like core-site.xml, hdfs-site.xml, mapred-site.xml based upon the job requirement.
- Set up and managing HA Name Node to avoid single point of failures in large clusters.
- Involved in Installing and configuring Kerberos for the authentication of users and Hadoop daemons.
- Monitoring workload, job performance, capacity planning.
- Experience in setting up tools like Ganglia for monitoring Hadoop cluster.
- Implemented Backup configurations and Recoveries from a Name Node failure.
- Worked with systems engineering team to plan and deploy new Hadoop environments and expand existing Hadoop clusters.
- Experience monitoring and troubleshooting issues with Linux memory, CPU, OS, storage and network.
- Hands on experience in analyzing Log files for Hadoop and eco system services and finding root cause.
- Experience in HDFS data storage and support for running map-reduce jobs.
- Experience in understanding the security requirements for Hadoop and integrating with Kerberos authentication infrastructure- KDC server setup, creating realm /domain, managing.
- Worked on Hadoop Stack, ETL TOOLS like Tableau and Security like Kerberos. User provisioning with LDAP and lot of other Big Data technologies for multiple use cases.
- Recovering from node failures and troubleshooting common Hadoop cluster issues.
- Trouble shooting many cloud related issues such as Data Node down, Network failure and data block missing.
- Scheduling and Managing Oozie Jobs to automate sequence of rotational activity.
- Managing Disk File Systems, Server Performance, Users Creation and Granting file access Permissions and RAID configurations.
- Utilization of dump and restore for file system backup and restoration.
- Monitoring systems and services through Ambari dashboard to make the clusters available for the business.
- Hand on experience on cluster up gradation and patch upgrade without any data loss and with proper backup plans.
- Modified reports and Talen ETL jobs based on the feedback from QA testers and Users in development and staging environments.
- Changing the configurations based on the requirements of the users for the better performance of the jobs.
- Provided security and authentication with ranger where ranger admin provides administration and user sync adds the new users to the cluster.
- Good troubleshooting skills on Hue, which provides GUI for developer's/business users for day to day activities.

**Environment:** Hadoop, Map Reduce, Yarn, Hive, HDFS, PIG, Sqoop, Solr, Oozie, Impala, Spark, Hortonworks, Flume, HBase, Zookeeper and Unix/Linux, Hue (Beeswax), AWS.

**Aug 2016 – Oct 2018**

**Hadoop Big data Administrator (System Engineer), USAA**

**San Antonio, TX**

**Responsibilities:**

- Built Cloudera Hadoop clusters.
- Responsible for cluster maintenance, adding and removing cluster nodes, cluster monitoring and troubleshooting, managing and reviewing data backups and Hadoop log files.
- Installed/Configured/Maintained Apache Hadoop clusters for application development and Hadoop tools like Hive, Pig, Spark, Kafka, Impala, Zookeeper, Hue and Sqoop using both Cloudera and Hortonworks.
- Involved in Upgrading the CDH and Cloudera Manager (5.8, 5.10 to 5.12).
- Enabled HA for Cloudera Manager.
- Enabled HA for Resource Manager, Name Node, HiveMetastore.
- Enabled Active Directory/LDAP for Cloudera Manager, Cloudera Navigator and Hue.
- Led the team in setting up Kerberos for Cloudera and Kafka cluster.
- Enabled load balancer for impala to distribute data load on all impala daemons across the cluster.
- Expertise in setting up in-memory layer such as Spark (1.6 and 2.x), impala and its maintenance like resolving out of memory issues, balancing load across daemons.
- Worked on Impala for interactive reporting for tableau, Pentaho.
- Worked on real time data processing with Spark.
- Provided guidance to users on re-writing their queries to improve performance and reduce cluster usage.
- Provided regular user and application support for highly complex issues involving multiple components such as Hive, Impala, Spark, Kafka, MapReduce.
- Handling data import from various data sources, performed transformations using Hive, Map Reduce, loaded data into HDFS and Extracted the data from Mysql into HDFS.
- Written scripts for disk monitoring and logs compression
- Involved in ongoing maintenance, support and improvement in Hadoop cluster.
- Created Kafka topics, provide ACLs to users and setting up rest mirror and mirror maker to transfer the data between two Kafka clusters.
- Helped the users to connect to Kerberized Hive from SQL Workbench and BI tools.
- Upgraded HDP and Ambari.

**Environment:** Hadoop Hdfs, Map reduce, Hive, Pig, Flume, Oozie, Sqoop, Eclipse, Horton works, Ambari.

**Aug 2015 – Jul 2016**

**Hadoop/Linux Administrator, AbbVie Inc.**

**Washington D.C.**

**Responsibilities:**

- Worked on Administration, maintenance and support of Red Hat Enterprise Linux (RHEL) servers.
- Executed user administration and maintenance tasks including creating users & groups.
- Upgrading packages and patching systems to maintain the productive environment using rpm and yum.
- Upgraded and maintained software packages on servers using RHEL Satellite and Repository servers.
- Resizing volumes to meet customer requirement, dealing with Volume Manager Performance issues.
- Writing Bash scripts for backup and automation.
- Install and configure DHCP, DNS (BIND, MS), web (Apache, IIS) and file servers on Linux servers.
- Periodic checks of production and development systems; CPU utilization, memory profiles, disk utilization, network connectivity, system log files, etc.
- Performed all System Administration tasks like installing packages, and patches.

- Troubleshooting various problems related to VM in initializing, replacing, mirroring, encapsulating and removing disk devices on various Production Boxes.
- Handled effectively Issues related to NFS, NIS, and LVM, configuration & maintenance of RAID Levels.
- Managing and troubleshooting user's login problems.
- Manage and monitor Active Directory services and group policies.
- Involved in development, user acceptance, and performance testing, production & disaster recovery server.
- Add swap space on disks as needed using Linux utilities.
- Experience in Installation of third-party tools using packages on Linux and Ubuntu.
- Upgraded and maintained software packages on servers using RHEL Satellite and Repository servers.
- Configured resources such as packages, services, files, directories, notifying users and groups, setting up Cron Scheduler.
- Performed all necessary day-to-day Git support, implemented, maintained the Branching, Build/Release strategies utilizing Git repositories.

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

**Aug 2012 – July 2014**

**Hadoop/Linux Administrator, Systematix InfoTech Pvt Ltd  
India**

**Responsibilities:**

- Worked on Multi Clustered environment and setting up Cloudera and Hortonworks Hadoop eco-System.
- Had good working experience on Hadoop architecture, HDFS, Map Reduce and other components in the Cloudera - Hadoop eco system.
- Installation, configuration and administration of Linux (RHEL 4/5) and Solaris 8/9/10 servers.
- Configured Red Hat Kick start server for installing multiple production servers.
- Involved in maintaining Hadoop cluster in development and test environment.
- Installing and partitioning disk drives. Creating, mounting and maintaining file systems to ensure access to system, application and user data.
- Performed upgrades, patches and bug fixes in HDP and CDH clusters.
- Maintenance and installation of RPM and YUM package installations and other server.
- Creating users, assigning groups and home directories, setting quota and permissions; administering file systems and recognizing file access problems.
- Experience in Managing and Scheduling Cron jobs such as enabling system logging, network logging of servers for maintenance, performance tuning and testing.
- Maintaining appropriate file and system security, monitoring and controlling system access, changing permission, ownership of files and directories, maintaining passwords, assigning special privileges to selected users and controlling file access.
- Extensive use of LVM, creating Volume Groups, Logical volumes.
- Performed various configurations which include networking and IPTables, resolving hostnames, SSH key less login.
- Build, configure, deploy, support, and maintain enterprise class servers and operating systems.
- Building Centos 5/6 Servers, Oracle Enterprise Linux 5 and RHEL (4/5) Servers from scratch.
- Organized the patch depots and act as POC for the patch related issues.
- Configuration & Installation of Red hat Linux 5/6, Cent OS 5 and Oracle Enterprise Linux 5 by using Kick Start to reduce the installation issues.
- Handled patch upgrades and firmware upgrades on and RHEL Servers, Oracle Enterprise Linux Servers.
- User and Group administration on RHEL Systems.
- Creation of various user profiles and environment variables to ensure security.
- Server hardening and security configurations as per the client specifications.

- A solid understanding of networking/distributed computing environment concepts, including principles of routing, bridging and switching, client/server programming, and the design of consistent network-wide file system layouts.
- Strong understanding of Network Infrastructure Environment.  
**Environment:** Red Hat Linux AIX, RHEL, Oracle 9i/10g, Samba, NT/2000 Server, VMware 2.x, Tomcat 5.x, Apache Server.

## **INTERESTS:**

- Blogging, Reading Articles
- Hiking and Travelling
- Listening Music

## **LANGUAGES**

- English
- Hindi

# Manikya Kasturi

- Morristown, NJ, USA

## Contact Information

- manikasturi5@gmail.com (Preferred)
- 2098131243 (Preferred)

## Work History

### Total Work Experience: 7 years

- **State farm | State farm**  
Nov 01, 2018 - No End Date
- **Hadoop Big data Administrator (System Engineer) | USAA**  
Aug 01, 2016 - Oct 01, 2018
- **Hadoop/Linux Administrator | AbbVie Inc.**  
Aug 01, 2015 - Jul 01, 2016
- **Hadoop/Linux Administrator | Systematix InfoTech Pvt Ltd**  
Aug 01, 2012 - Jul 01, 2014

## Skills

- **scripting** | 7yrs | 2019
- **troubleshooting** | 7yrs | 2019
- **apache kafka** | 7yrs | 2019
- **apache sqoop** | 7yrs | 2019
- **hdfs** | 7yrs | 2019
- **database** | 6yrs | 2019
- **operating systems** | 6yrs | 2019
- **qa** | 6yrs | 2019
- **rbdbms** | 6yrs | 2019

- **reporting** | 6yrs | 2019
- **apache hadoop** | 6yrs | 2019
- **distribution** | 6yrs | 2019
- **installation** | 6yrs | 2019
- **software** | 6yrs | 2019
- **mapreduce** | 5yrs | 2019
- **backup** | 5yrs | 2019
- **configuration** | 5yrs | 2019
- **hortonworks** | 5yrs | 2019
- **monitoring** | 5yrs | 2019
- **linux** | 4yrs | 2019
- **active directory** | 4yrs | 2019
- **apache pig** | 4yrs | 2019
- **cluster** | 4yrs | 2019
- **file systems** | 4yrs | 2019
- **apache flume** | 4yrs | 2019
- **apache hbase** | 4yrs | 2019
- **apache hive** | 4yrs | 2019
- **apache oozie** | 4yrs | 2019
- **apache spark** | 4yrs | 2019
- **network** | 4yrs | 2019
- **poc** | 4yrs | 2019
- **big data** | 3yrs | 2019
- **kerberos** | 3yrs | 2019
- **apache zookeeper** | 3yrs | 2019
- **cloudera impala** | 3yrs | 2019
- **technical support** | 5yrs | 2018
- **cloudera** | 4yrs | 2018
- **cdh** | 3yrs | 2018
- **nosql** | 5yrs | 2016
- **lvm** | 4yrs | 2016
- **cron** | 4yrs | 2016
- **red hat enterprise linux** | 4yrs | 2016
- **rpm** | 4yrs | 2016
- **apache**
- **hadoop**
- **hbase**
- **hive**

## Work Preferences

- Likely to Switch: Most Likely
- Willing to Relocate: Yes
- Work Authorization:

- US
- Work Documents:
  - Have H1
- Security Clearance: No
- Third Party: Yes
- Employment Type:
  - Contract - Corp-to-Corp

## Profile Sources

- Dice:  
<https://www.dice.com/employer/talent/profile/935604294696d71c5cce89e7270a7dbcd1026984>