|  |  |
| --- | --- |
| **DOCUMENT RULES:** | |
| **Documentation Name:** | **Hadoop Cluster** |
| **Task name & column name should be written:** | **Bold (CTRL+B)** |
| **Commands should be written in the after # sign:** | *Italic (CTRL+I) #hostname* |
| **Output photo should be cropped or compressed:**  **Photo could be more than one:**  **If you need extra lines, add the line next after it:** | ***Description photo should be with title bar (CTRL + I + B)*** |
| **All other text should be written:** | Standard |
| **Font name and text size:** | Calibri and 9 |
| **Group name:** | 652.20E |
| **E-mail:** | 652.20e@gmail.com |

|  |  |  |
| --- | --- | --- |
| **#** | **Task names** | **Command steps and outputs** |
|  | **Check status of firewall and take screenshot of the CLI output.**  **If firewall is not installed left it as have.**  **Give permanent SSH access from.**  **SSH virtual port is 22. If SSH server is not installed, please install it.**  **To be check connectivity use commands.**  **Update and upgrade the system and application.** | ***#****sudo apt-get update*  ***#****sudo apt-get upgrade*  ***#****ping x.x.x.x*  ***#****telnet x.x.x.x22*  ***#****sudo ufw status* |
|  | ****Install ssh and pdsh**** | ***#****sudo apt install shh*  ***#****sudo apt install pdsh* |
|  | **Set pdsh environment to ssh** | ***#****sudo nano .bashrc*  ***#****export PDSH\_RCMD\_TYPE=ssh* |
|  | ****Generate a SSH key**** | ***#****ssh-keygen -t rsa -P ""*  ***#****cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys* |
|  | **Installing java** | ***#****sudo apt install openjdk-8-jdk* |
|  | **Installing Hadoop, unzip and move folder** | ***#****sudo wget -P ~* [*https://mirrors.sonic.net/apache/hadoop/common/hadoop-3.2.1/hadoop-3.2.1.tar.gz*](https://mirrors.sonic.net/apache/hadoop/common/hadoop-3.2.1/hadoop-3.2.1.tar.gz)  ***#****tar xzf hadoop-3.2.1.tar.gz*  ***#****mv hadoop-3.2.1 hadoop* |
|  | **Export java** | ***#****sudo nano ~/hadoop/etc/hadoop/hadoop-env.sh*  ***#*** *~ export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64/* |
|  | Move the hadoop directory to our user local file | ***#****sudo mv hadoop /usr/local/hadoop* |
|  | Set up hadoop path | ***#****sudo nano /etc/environment*  ***#****~PATH="/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin:/usr/games:/usr/local/games:/usr/local/hadoop/bin:/usr/local/hadoop/sbin"*  *JAVA\_HOME="/usr/lib/jvm/java-8-openjdk-amd64/jre"* |
|  | Create a specific user for Hadoop | ***#****sudo adduser h-user* |
|  | **Give this user permissions to work within hadoop’s folder** | ***#****sudo usermod -aG hadoopuser h-user*  ***#****sudo chown h-user:root -R /usr/local/hadoop/*  ***#****sudo chmod g+rwx -R /usr/local/hadoop/*  ***#****sudo adduser h-user sudo* |
|  | Change hostnames | ***#****sudo nano /etc/hosts* |
|  | Configure Hadoop Service Port **(only on primary)** | ***#****sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml*  ***#****~<property>*  *<name>fs.defaultFS</name>*  *<value>hdfs://h-primary:9000</value>*  *</property>* |
|  | Configuration of HDFS system **(only on primary)** | ***#****sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml*  ***#****~<property>*  *<name>dfs.namenode.name.dir</name><value>/usr/local/hadoop/data/nameNode</value>*  *</property>*  *<property>*  *<name>dfs.datanode.data.dir</name><value>/usr/local/hadoop/data/dataNode</value>*  *</property>*  *<property>*  *<name>dfs.replication</name>*  *<value>2</value>*  *</property>* |
|  | Identify the workers **(only on primary)** | ***#****sudo nano /usr/local/hadoop/etc/hadoop/workers*  ***#****~scp /usr/local/hadoop/etc/hadoop/\* h-secondary1:/usr/local/hadoop/etc/hadoop/*  *scp /usr/local/hadoop/etc/hadoop/\* h-secondary2:/usr/local/hadoop/etc/hadoop/* |
|  | Formatting and Starting HDFS system (only primary) | ***#****source /etc/environment*  ***#****hdfs namenode -format* |
|  | **Update changes and start service** | ***#****source ~/.bashrc*  ***#****start-dfs.sh* |
|  | **Start Yarn service** | ***#****start-yarn.sh* |