Arth Task 7.1 : Elasticity Task : Increase the storage of a data-node to increase the overall storage: Integrating LVM with Hadoop

First we create ah Hadoop Cluster with 1 master and 1 Datanode

```
Datanodes available: 1 (1 total, 0 dead)

Name: 65.0.71.177:50010
Decommission Status : Normal
Configured Capacity: 5217320960 (4.86 GB)
DFS Used: 8192 (8 KB)
Non DFS Used: 306167808 (291.98 MB)
DFS Remaining: 4911144960(4.57 GB)
DFS Used%: 0%
DFS Remaining%: 94.13%
Last contact: Sat Mar 13 07:22:17 UTC 2021
```

Before Increasing or decreasing the memory , we check the total number of partitions in the OS using : **fdisk -l**

```
[root@ip-172-31-13-235 ec2-user]# fdisk -l
Disk /dev/xvda: 10 GiB, 10737418240 bytes, 20971520 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: gpt
Disk identifier: 246B752E-8CB4-41E7-B9B1-365A93ACF890

Device      Start       End   Sectors Size Type
/dev/xvda1  2048       4095      2048   1M BIOS boot
/dev/xvda2  4096 20971486 20967391  10G Linux filesystem
[root@ip-172-31-13-235 ec2-user]#
```

Next we add a Secondary storage to the OS : Since am using AWS , I use EBS to add a secondary storage

Attach Volume   ✕

Volume ⓘ	vol-0eeba986d16a3089c in ap-south-1b	
Instance ⓘ	i-01e5a4323b7fa0dd1	in ap-south-1b
Device ⓘ	/dev/sdf	
	Linux Devices: /dev/sdf through /dev/sdp	

Note: Newer Linux kernels may rename your devices to /dev/xvdf through /dev/xvdp internally, even when the device name entered here (and shown in the details) is /dev/sdf through /dev/sdp.

Cancel   **Attach**

```
[root@MiWiFi-R3L-srv ~]# fdisk -l
Disk /dev/sda: 20 GiB, 21523628032 bytes, 42038336 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: dos
Disk identifier: 0x8628e552

Device     Boot    Start       End  Sectors Size Id Type
/dev/sda1   *        2048   2099199  2097152   1G 83 Linux
/dev/sda2         2099200  42037247 39938048  19G 8e Linux LVM


Disk /dev/mapper/rhel-root: 17 GiB, 18291359744 bytes, 35725312 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes


Disk /dev/mapper/rhel-swap: 2 GiB, 2155872256 bytes, 4210688 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
[root@MiWiFi-R3L-srv ~]#
```

Now we have to create a physical volume for the storage using command: **pvcreate /dev/xvdf** and to confirm we use command **: pvdisplay**

```
[root@ip-172-31-13-235 hadoop]# pvcreate /dev/xvdf
  Physical volume "/dev/xvdf" successfully created.
[root@ip-172-31-13-235 hadoop]# pvdisplay
  "/dev/xvdf" is a new physical volume of "10.00 GiB"
  --- NEW Physical volume ---
  PV Name                   /dev/xvdf
  VG Name
  PV Size                   10.00 GiB
  Allocatable               NO
  PE Size                   0
  Total PE                  0
  Free PE                   0
  Allocated PE              0
  PV UUID                   vYZ5lb-bma1-nMKq-1Lij-t6Zd-u44Q-mbirPJ
```

After creating Physical volume , we need to create a volume group using command : **vgcreate [Group-name] /dev/xvdf** and to confirm we use command **:vgdisplay**

```
[root@ip-172-31-13-235 hadoop]# vgcreate Rahul /dev/xvdf
  Volume group "Rahul" successfully created
[root@ip-172-31-13-235 hadoop]# vgdisplay
  --- Volume group ---
  VG Name                   Rahul
  System ID
  Format                    lvm2
  Metadata Areas            1
  Metadata Sequence No      1
  VG Access                 read/write
  VG Status                 resizable
  MAX LV                    0
  Cur LV                    0
  Open LV                   0
  Max PV                    0
  Cur PV                    1
  Act PV                    1
  VG Size                   <10.00 GiB
  PE Size                   4.00 MiB
  Total PE                  2559
  Alloc PE / Size           0 / 0
  Free  PE / Size           2559 / <10.00 GiB
  VG UUID                   WWfPdM-nLx0-P6kd-d5hN-pTLh-mwbU-hAnFzJ
```

Finally creating a Logical Volume using the previous created volume group using command : **lvcreate — size [Storage amount] — name [group-name] [physical volume]**

```
[root@ip-172-31-13-235 hadoop]# lvcreate --size 5G --name LVM1 Rahul
  Logical volume "LVM1" created.
[root@ip-172-31-13-235 hadoop]# lvdisplay
  --- Logical volume ---
  LV Path                /dev/Rahul/LVM1
  LV Name                LVM1
  VG Name                Rahul
  LV UUID                CD8cud-wuf7-uwCp-09YL-EURE-tbKh-GcCIqd
  LV Write Access        read/write
  LV Creation host, time ip-172-31-13-235.ap-south-1.compute.internal,
3 07:18:37 +0000
  LV Status              available
  # open                 0
  LV Size                5.00 GiB
  Current LE             1280
  Segments               1
  Allocation             inherit
  Read ahead sectors     auto
  - currently set to     8192
  Block device           253:0
```

We need to format the volume before mounting using command :

mkfs.ext4 /dev/[Groupname]/[partition name]

```
Disk /dev/mapper/Rahul-LVM1: 5 GiB, 5368709120 bytes, 10485760 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
[root@ip-172-31-13-235 hadoop]# mkfs.ext4 /dev/Rahul/LVM1
mke2fs 1.45.6 (20-Mar-2020)
Creating filesystem with 1310720 4k blocks and 327680 inodes
Filesystem UUID: ea33cfda-0f30-4429-8f97-230429306fef
Superblock backups stored on blocks:
        32768, 98304, 163840, 229376, 294912, 819200, 884736

Allocating group tables: done
Writing inode tables: done
Creating journal (16384 blocks): done
Writing superblocks and filesystem accounting information: done
```

Finally , we mount the volume using command :

mount /dev/[Partition-name]/[Group-name] /storage-name

```
[root@ip-172-31-13-235 etc]# mount /dev/Rahul/LVM1 /dn
[root@ip-172-31-13-235 etc]# df -hT
Filesystem              Type      Size  Used Avail Use% Mounted on
devtmpfs                devtmpfs  378M     0  378M   0% /dev
tmpfs                   tmpfs     403M     0  403M   0% /dev/shm
tmpfs                   tmpfs     403M   16M  388M   4% /run
tmpfs                   tmpfs     403M     0  403M   0% /sys/fs/cgroup
/dev/xvda2              xfs        10G  1.8G  8.3G  18% /
tmpfs                   tmpfs      81M     0   81M   0% /run/user/1000
/dev/mapper/Rahul-LVM1  ext4      4.9G   20M  4.6G   1% /dn
```

Thus the storage has been increased , to confirm we can use command :

hadoop dfsadmin -report

```
[root@ip-172-31-14-252 hadoop]# hadoop dfsadmin -report
Configured Capacity: 10724814848 (9.99 GB)
Present Capacity: 8818454528 (8.21 GB)
DFS Remaining: 8818446336 (8.21 GB)
DFS Used: 8192 (8 KB)
DFS Used%: 0%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0

-------------------------------------------------
Datanodes available: 1 (1 total, 0 dead)

Name: 65.0.71.177:50010
Decommission Status : Normal
Configured Capacity: 10724814848 (9.99 GB)
DFS Used: 8192 (8 KB)
Non DFS Used: 1906360320 (1.78 GB)
DFS Remaining: 8818446336(8.21 GB)
DFS Used%: 0%
DFS Remaining%: 82.22%
Last contact: Sat Mar 13 06:50:46 UTC 2021
```

Thus it is confirmed that the storage has been increased by 5 GB.

This whole process can be automated using python scripting :

GITHUB URL : https://github.com/Rahul1498/ARTH-Tasks/blob/main/LVMautomation.py