REPORT

Overview

You are employed at a company which creates dashboard/reports on a daily basis. There are two main stages that occur each day to produce the dashboards:

- Ingestion.
- Filtering and reporting.

In the first stage, the company downloads csv files daily. While in the second stage, these csv files are filtered (cleaned for analyses) and analysed to produce the dashboards. On average, the filtering stage filters roughly 25% of the ingested data on a daily basis.

Problem Statement: Currently the company are facing issues regarding csv files storage whereas they are finding it difficult to scale their storage device since they have to manually increase the size every time they need more space. They are currently storing everything on a single partition. Extending this partition is often difficult and tiresome, having to unmount the partition and copy/move data to extend the partition.
Your job is to reorganize their storage system, making it more robust, extensible, automated and overall easier to manage and scale by performing the following:

1. Storage system design using logical volume managers.
2. Automation and auto-scaling.
3. SELinux employment.

Pre-requisites

- The system you are working on should have a 8 GB external storage disk (check the storage supplementary notes pdf posted on the cms to know how to add a disk on a virtual machine in Virtual Box)

Done by:

Omar Ahmed

52-23411

Lab day: Thursday

Screenshot(s):

a) lsblk-fs

```
[root@localhost ~]# lsblk -fs
NAME                FSTYPE      LABEL UUID                                   MOUNTPOINT
sda1                xfs               d695fa3d-7f9a-4e7a-98ae-031e672b4d64   /boot
└─sda
sr0
rhel-root           xfs               28a81415-ffc1-4697-8716-4aa7005f809a   /
└─sda2              LVM2_member       1hXCiK-XmWX-9M04-qOun-An9f-Lu2v-uDi6u4
  └─sda
rhel-swap           swap              4b2a8329-e04f-4e51-9fe2-fc9b087aa8cd   [SWAP]
└─sda2              LVM2_member       1hXCiK-XmWX-9M04-qOun-An9f-Lu2v-uDi6u4
  └─sda
vg_data-lv_ingestion xfs              8afa2514-044f-4bad-8aee-219caad2e45a   /root/ingestion
└─sdb1              LVM2_member       9uQ3s5-wQ5H-cEKv-WNNM-kgxo-bJBG-VvWRGJ
  └─sdb
vg_data-lv_analyses  xfs              a2834cd9-eda7-469a-9b69-35a823981ae1   /root/analyses
└─sdb1              LVM2_member       9uQ3s5-wQ5H-cEKv-WNNM-kgxo-bJBG-VvWRGJ
  └─sdb
[root@localhost ~]#
```

b) cat /etc/fstab

```
[root@localhost ~]# cat /etc/fstab

#
# /etc/fstab
# Created by anaconda on Wed Aug 19 10:06:24 2020
#
# Accessible filesystems, by reference, are maintained under '/dev/disk/'.
# See man pages fstab(5), findfs(8), mount(8) and/or blkid(8) for more info.
#
# After editing this file, run 'systemctl daemon-reload' to update systemd
# units generated from this file.
#
/dev/mapper/rhel-root   /                       xfs     defaults        0 0
UUID=d695fa3d-7f9a-4e7a-98ae-031e672b4d64 /boot                   xfs     defaults        0 0
/dev/mapper/rhel-swap   swap                    swap    defaults        0 0
UUID=8afa2514-044f-4bad-8aee-219caad2e45a /root/ingestion xfs defaults 0 0
UUID=a2834cd9-eda7-469a-9b69-35a823981ae1 /root/analyses xfs defaults 0 0
[root@localhost ~]#
```

c) findmnt-verify

```
[root@localhost ~]# findmnt --verify
/
   [W] recommended root FS passno is 1 (current is 0)

0 parse errors, 0 errors, 1 warning
[root@localhost ~]#
```

d) pvdisplay , vgdisplay , lvdisplay

```
[root@localhost ~]# pvdisplay /dev/sdb1
  --- Physical volume ---
  PV Name               /dev/sdb1
  VG Name               vg_data
  PV Size               <8.00 GiB / not usable 2.00 MiB
  Allocatable           yes
  PE Size               4.00 MiB
  Total PE              2047
  Free PE               511
  Allocated PE          1536
  PV UUID               9uQ3s5-wQ5H-cEKv-WNNM-kgxo-bJBG-VvWRGJ

[root@localhost ~]#
```

```
[root@localhost ~]# vgdisplay vg_data
  --- Volume group ---
  VG Name               vg_data
  System ID
  Format                lvm2
  Metadata Areas        1
  Metadata Sequence No  3
  VG Access             read/write
  VG Status             resizable
  MAX LV                0
  Cur LV                2
  Open LV               2
  Max PV                0
  Cur PV                1
  Act PV                1
  VG Size               <8.00 GiB
  PE Size               4.00 MiB
  Total PE              2047
  Alloc PE / Size       1536 / 6.00 GiB
  Free  PE / Size       511 / <2.00 GiB
  VG UUID               rtrUmT-BNZc-SUYr-93pj-gaUg-ewJ6-iafBI3

[root@localhost ~]#
```

```
[root@localhost ~]# lvdisplay vg_data/lv_ingestion
  --- Logical volume ---
  LV Path                /dev/vg_data/lv_ingestion
  LV Name                lv_ingestion
  VG Name                vg_data
  LV UUID                dZlCH9-kdvt-RsTj-hQbc-sSde-FCZf-BXGnYR
  LV Write Access        read/write
  LV Creation host, time localhost.localdomain, 2025-05-09 19:50:42 +0200
  LV Status              available
  # open                 1
  LV Size                4.00 GiB
  Current LE             1024
  Segments               1
  Allocation             inherit
  Read ahead sectors     auto
  - currently set to     8192
  Block device           253:2

[root@localhost ~]#
```

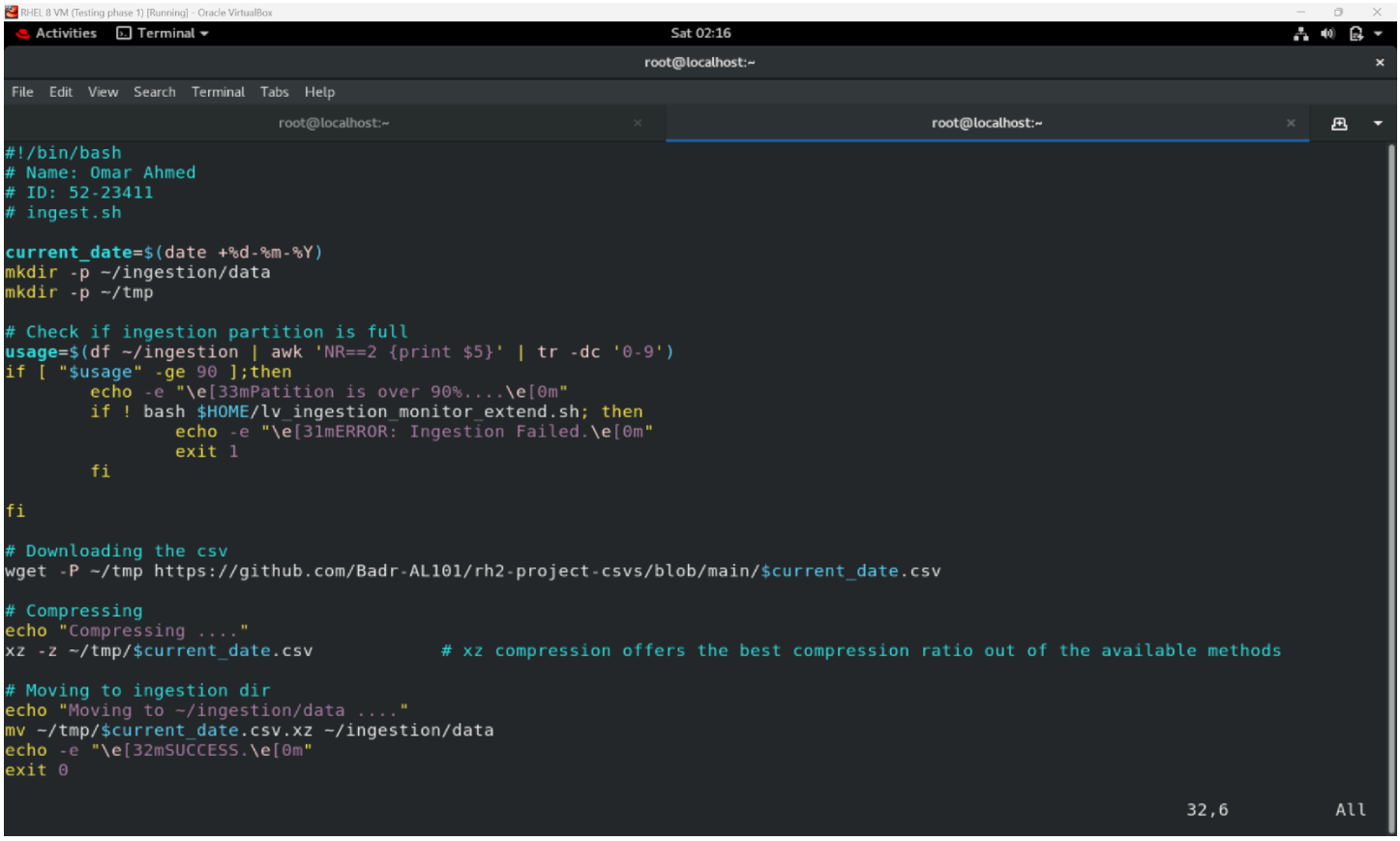
```
[root@localhost ~]# lvdisplay vg_data/lv_analyses
  --- Logical volume ---
  LV Path                /dev/vg_data/lv_analyses
  LV Name                lv_analyses
  VG Name                vg_data
  LV UUID                eGujss-R1y8-0WRe-dIJ1-CfJ8-mufH-Nl0DxU
  LV Write Access        read/write
  LV Creation host, time localhost.localdomain, 2025-05-09 19:50:53 +0200
  LV Status              available
  # open                 1
  LV Size                2.00 GiB
  Current LE             512
  Segments               1
  Allocation             inherit
  Read ahead sectors     auto
  - currently set to     8192
  Block device           253:3

[root@localhost ~]#
```

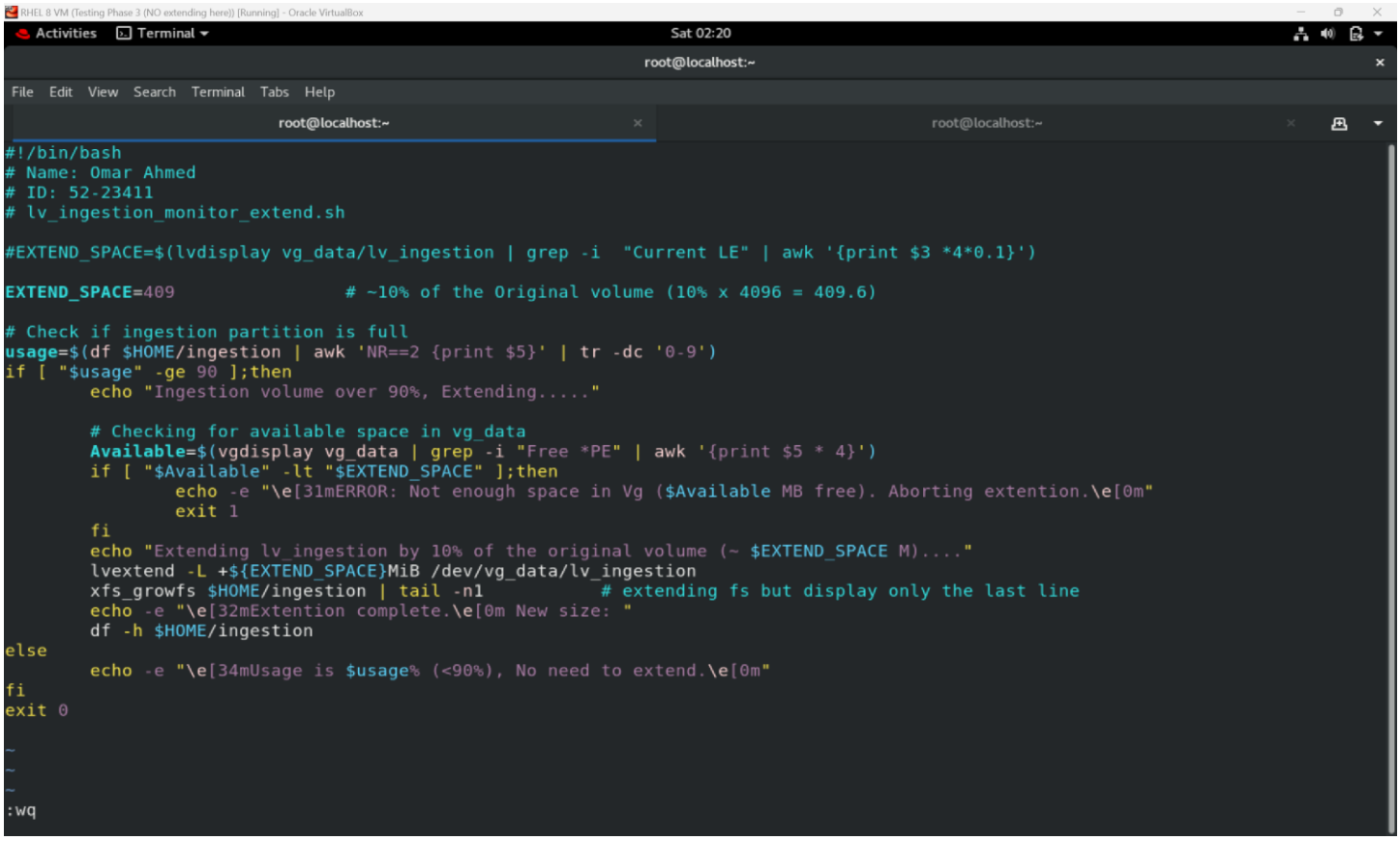
e) Screenshot of your cron table

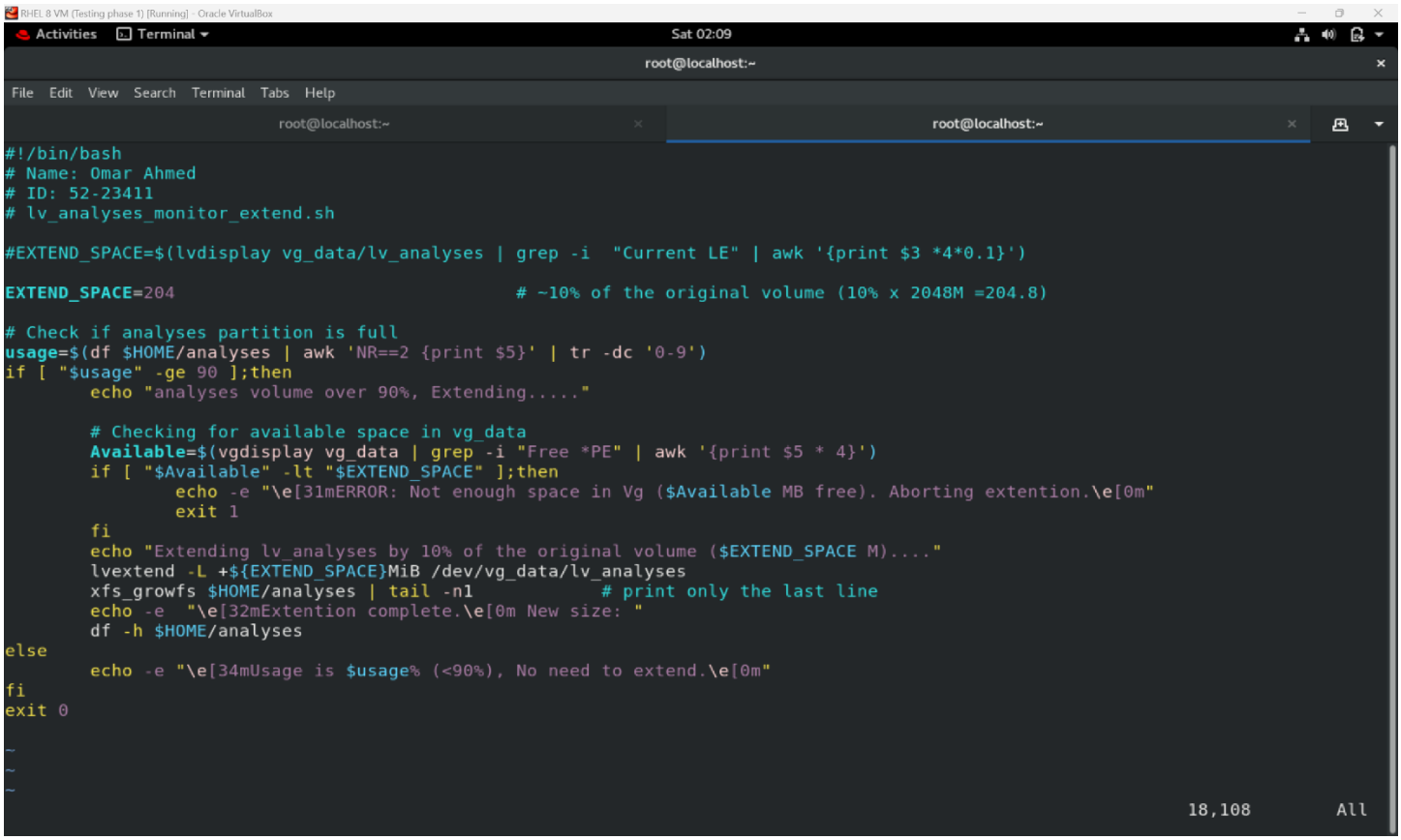
- (10:00 am for the monitoring and 11:00 am for the ingestion at a daily basis)

```
[root@localhost ~]# crontab -l
0 11 * * * /root/ingest.sh
0 10 * * * /root/lv_ingestion_monitor_extend.sh
0 10 * * * /root/lv_analyses_monitor_extend.sh
[root@localhost ~]#
```

f) Screenshot of all your scripts IMPORTANT-YOUR name and id must be included at the start of each script.







g) semanage fcontext-l-C

```
[root@localhost ~]# semanage fcontext -l -C
SELinux fcontext                                    type               Context

/root/analyses/reports(/.*)?                        all files          system_u:object_r:httpd_sys_content_t:s0
[root@localhost ~]#
```

h) ls-Z /analyses/reports

- create demo test files to test the correct context

```
[root@localhost ~]# ls -aZ analyses/reports/
    system_u:object_r:httpd_sys_content_t:s0 .                 unconfined_u:object_r:httpd_sys_content_t:s0 test_report2.docx
            system_u:object_r:unlabeled_t:s0 ..                unconfined_u:object_r:httpd_sys_content_t:s0 test_report3.docx
unconfined_u:object_r:httpd_sys_content_t:s0 test_report1.docx  unconfined_u:object_r:httpd_sys_content_t:s0 test_report4.docx
[root@localhost ~]#
```

i) ls-al /ingestion/data

```
[root@localhost ~]# ls -al ingestion/data/
total 56
drwxr-xr-x. 2 root root    56 May 12 21:26 .
drwxr-xr-x. 3 root root    18 May 10 01:03 ..
-rw-r--r--. 1 root root 27984 May 10 02:55 10-05-2025.csv.xz
-rw-r--r--. 1 root root 28016 May 12 21:26 12-05-2025.csv.xz
[root@localhost ~]#
```

Testing:

a) Scheduling the tasks to be done @ 10:25pm and 10:26 pm (For testing purposes)

```
[root@localhost ~]# grep "\May 12"  /var/log/cron
May 12 21:18:25 localhost crontab[15656]: (root) BEGIN EDIT (root)
May 12 21:19:09 localhost crontab[15656]: (root) REPLACE (root)
May 12 21:19:09 localhost crontab[15656]: (root) END EDIT (root)
May 12 21:18:25 localhost crontab[15656]: (root) BEGIN EDIT (root)
May 12 21:19:09 localhost crontab[15656]: (root) REPLACE (root)
May 12 21:19:09 localhost crontab[15656]: (root) END EDIT (root)
May 12 21:24:12 localhost crontab[16289]: (root) BEGIN EDIT (root)
May 12 21:24:45 localhost crontab[16289]: (root) REPLACE (root)
May 12 21:24:45 localhost crontab[16289]: (root) END EDIT (root)
May 12 21:25:01 localhost crond[888]: (root) RELOAD (/var/spool/cron/root)
May 12 21:25:01 localhost CROND[16327]: (root) CMD (/root/lv_analyses_monitor_extend.sh)
May 12 21:25:01 localhost CROND[16329]: (root) CMD (/root/lv_ingestion_monitor_extend.sh)
May 12 21:25:01 localhost CROND[16323]: (root) CMDOUT (#033[34mUsage is 3% (<90%), No need to extend.#033[0m)
May 12 21:25:01 localhost CROND[16324]: (root) CMDOUT (#033[34mUsage is 2% (<90%), No need to extend.#033[0m)
May 12 21:26:01 localhost CROND[16371]: (root) CMD (/root/ingest.sh)
May 12 21:26:01 localhost CROND[16369]: (root) CMDOUT (--2025-05-12 21:26:01--  https://github.com/Badr-AL101/rh2-project-csvs/blob/main/12-05-2025.csv)
May 12 21:26:01 localhost CROND[16369]: (root) CMDOUT (Resolving github.com (github.com)... 140.82.121.3)
May 12 21:26:02 localhost CROND[16369]: (root) CMDOUT (Connecting to github.com (github.com)|140.82.121.3|:443... connected.)
May 12 21:26:02 localhost CROND[16369]: (root) CMDOUT (HTTP request sent, awaiting response... 200 OK)
May 12 21:26:02 localhost CROND[16369]: (root) CMDOUT (Length: unspecified [text/html])
May 12 21:26:02 localhost CROND[16369]: (root) CMDOUT (Saving to: '/root/tmp/12-05-2025.csv')
May 12 21:26:02 localhost CROND[16369]: (root) CMDOUT ()
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (     0K .......... .......... .......... .......... ..........  235K)
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (    50K .......... .......... .......... .......... ..........  466K)
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (   100K .......... .......... .......... .......... ..........  387K)
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (   150K .......... .......... ..                                776K=0.5s)
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT ()
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (2025-05-12 21:26:03 (360 KB/s) - '/root/tmp/12-05-2025.csv' saved [176473])
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT ()
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (Compressing ....)
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (Moving to ~/ingestion/data ....)
May 12 21:26:03 localhost CROND[16369]: (root) CMDOUT (#033[32mSUCCESS.#033[0m)
[root@localhost ~]# ls ingestion/data/
10-05-2025.csv.xz  12-05-2025.csv.xz
[root@localhost ~]#
```

b) Trying to extend when there is no need

```
[root@localhost ~]# ./lv_ingestion_monitor_extend.sh
Usage is 2% (<90%), No need to extend.
[root@localhost ~]# ./lv_analyses_monitor_extend.sh
Usage is 3% (<90%), No need to extend.
[root@localhost ~]#
```

c) Modifying the scripts to extend when >1%

```
#!/bin/bash
# Name: Omar Ahmed
# ID: 52-23411
# lv_ingestion_monitor_extend.sh

#EXTEND_SPACE=$(lvdisplay vg_data/lv_ingestion | grep -i  "Current LE" | awk '{print $3 *4*0.1}')

EXTEND_SPACE=409                # ~10% of the Original volume (10% x 4096 = 409.6)

# Check if ingestion partition is full
usage=$(df $HOME/ingestion | awk 'NR==2 {print $5}' | tr -dc '0-9')
if [ "$usage" -ge 1 ];then
        echo "Ingestion volume over 1%, Extending....."

        # Checking for available space in vg_data
        Available=$(vgdisplay vg_data | grep -i "Free *PE" | awk '{print $5 * 4}')
        if [ "$Available" -lt "$EXTEND_SPACE" ];then
                echo -e "\[31mERROR: Not enough space in Vg ($Available MB free). Aborting extention.\e[0m"
                exit 1
        fi
        echo "Extending lv_ingestion by 10% of the original volume (~ $EXTEND_SPACE M)...."
        lvextend -L +${EXTEND_SPACE}MiB /dev/vg_data/lv_ingestion
        xfs_growfs $HOME/ingestion | tail -n1           # extending fs but display only the last line
        echo -e "\e[32mExtention complete.\e[0m New size: "
        df -h $HOME/ingestion
else
        echo -e "\e[34mUsage is $usage% (<90%), No need to extend.\e[0m"
fi
exit 0
```

a. Trying to extend again (4 times)

```
[root@localhost ~]# ./lv_ingestion_monitor_extend.sh
Ingestion volume over 1%, Extending.....
Extending lv_ingestion by 10% of the original volume (~ 409 M)....
  Rounding size to boundary between physical extents: 412.00 MiB.
  Size of logical volume vg_data/lv_ingestion changed from 4.00 GiB (1024 extents) to 4.40 GiB (1127 extents).
  Logical volume vg_data/lv_ingestion successfully resized.
data blocks changed from 1048576 to 1154048
Extention complete. New size:
Filesystem                      Size  Used Avail Use% Mounted on
/dev/mapper/vg_data-lv_ingestion  4.4G   64M  4.4G   2% /root/ingestion
[root@localhost ~]# ./lv_ingestion_monitor_extend.sh
Ingestion volume over 1%, Extending.....
Extending lv_ingestion by 10% of the original volume (~ 409 M)....
  Rounding size to boundary between physical extents: 412.00 MiB.
  Size of logical volume vg_data/lv_ingestion changed from 4.40 GiB (1127 extents) to 4.80 GiB (1230 extents).
  Logical volume vg_data/lv_ingestion successfully resized.
data blocks changed from 1154048 to 1259520
Extention complete. New size:
Filesystem                      Size  Used Avail Use% Mounted on
/dev/mapper/vg_data-lv_ingestion  4.8G   67M  4.8G   2% /root/ingestion
[root@localhost ~]# ./lv_ingestion_monitor_extend.sh
Ingestion volume over 1%, Extending.....
Extending lv_ingestion by 10% of the original volume (~ 409 M)....
  Rounding size to boundary between physical extents: 412.00 MiB.
  Size of logical volume vg_data/lv_ingestion changed from 4.80 GiB (1230 extents) to <5.21 GiB (1333 extents).
  Logical volume vg_data/lv_ingestion successfully resized.
data blocks changed from 1259520 to 1364992
Extention complete. New size:
Filesystem                      Size  Used Avail Use% Mounted on
/dev/mapper/vg_data-lv_ingestion  5.2G   70M  5.2G   2% /root/ingestion
[root@localhost ~]# ./lv_ingestion_monitor_extend.sh
Ingestion volume over 1%, Extending.....
Extending lv_ingestion by 10% of the original volume (~ 409 M)....
  Rounding size to boundary between physical extents: 412.00 MiB.
  Size of logical volume vg_data/lv_ingestion changed from <5.21 GiB (1333 extents) to <5.61 GiB (1436 extents).
  Logical volume vg_data/lv_ingestion successfully resized.
data blocks changed from 1364992 to 1470464
Extention complete. New size:
Filesystem                      Size  Used Avail Use% Mounted on
/dev/mapper/vg_data-lv_ingestion  5.6G   73M  5.6G   2% /root/ingestion
```

b. Trying to extend again when there is no space left in Vg_data

```
[root@localhost ~]# ./lv_ingestion_monitor_extend.sh
Ingestion volume over 1%, Extending.....
ERROR: Not enough space in Vg (396 MB free). Aborting extention.
[root@localhost ~]# vgdisplay vg_data
  --- Volume group ---
  VG Name               vg_data
  System ID
  Format                lvm2
  Metadata Areas        1
  Metadata Sequence No  7
  VG Access             read/write
  VG Status             resizable
  MAX LV                0
  Cur LV                2
  Open LV               2
  Max PV                0
  Cur PV                1
  Act PV                1
  VG Size               <8.00 GiB
  PE Size               4.00 MiB
  Total PE              2047
  Alloc PE / Size       1948 / <7.61 GiB
  Free  PE / Size       99 / 396.00 MiB
  VG UUID               rtrUmT-BNZc-SUYr-93pj-gaUg-ewJ6-iafBI3
```