Cron Issue on SysLog Servers while Purging and Compressing logs

Syslog server log cleanup and compress cron's are currently not functioning as it was intended. This has caused the "/var/log/syslog" dataStore to fill up near capacity on most of the syslog servers at various DataCenters.

<u>Current Crons for deleting and compressing on the syslog servers</u>:

```
#cron to delete 90 days under /var/log/syslog

30 1 * * 0 (/bin/date && /usr/bin/find /var/log/syslog/ -type f -mtime +90 -exec ls -1 \; | awk -F: '/text/{print $1}'|while read FILE;do /bin/echo -n "Deleting $FILE... ";/bin/rm $FILE;/bin/echo "DONE";done) >> /var/log/cron_deleted_log

#cron to compress logs under /var/log/syslog

30 1 * * 0 (/bin/date && /usr/bin/find /var/log/syslog/ -type f -mtime +1 -exec /usr/bin/file {} \; | awk -F: '/text/{print $1}'|while read FILE;do /bin/echo -n "Compressing $FILE... ";/usr/bin/gzip -9 $FILE;/bin/echo "DONE";done) >> /var/log/cron_compressed_log

mo-2d57b9318[pc18splunklog01]:/var/log/syslog #
```

<u>Upon debugging,</u> why the cleanup and compression is not happening, found the "find" command from the cron would fail as below, when run from command line.

```
/usr/bin/find /var/log/syslog/ -type f -mtime +90 -exec ls -1 \;

/usr/bin/find: WARNING: Hard link count is wrong for `/var/log/syslog/' (saw only st_nlink=85 but we already saw 83 subdirectories): this may be a bug in your file system driver.  Automatically turning on find's -noleaf option.  Earlier results may have failed to include directories that should have been searched.
```

Drawbacks with current cron, for both deletion and compression of old logs:

a) Find, command hitting the above bug.

b) No logging of the files that were removed.

c) No validation to check, if the NFS mount (/var/log/syslog) is mounted for cleanup & compress.

To address the above issues:
Came up with new scripts (<u>purge90daysLogV1.bash</u> and <u>compressLogV2.bash</u>):

a) To overcome the find bug, additional options to the find command are used.

b) Audit log for files that were removed (90+ days older) with datestamp for each run (eg: p90log-20180605).

c) Audit log for compressed files (5 days older) with datestamp for each run (eg: CompressLog-20180605).

Manjesh, Platform OSS team

d) End of each audit log will have stats of the disk usage, before the script was run and after the run.

Eg.

```
mo-2d57b9318[pc18splunklog01]:/var/log/syslog # tail -4 purgelog/p90log-20180605

Syslog Before Clean Up : Filesystem                    Size  Used Avail Use% Mounted on

10.240.12.91:/vol1_splunklogs  3.0T  188G  2.9T   7% /var/log/syslog

Syslog After Clean Up : Filesystem                     Size  Used Avail Use% Mounted on

10.240.12.91:/vol1_splunklogs  3.0T   69G  3.0T   3% /var/log/syslog

mo-2d57b9318[pc18splunklog01]:/var/log/syslog #
```

Progress with the new scripts:

Except DC17, the new scripts have been run and tested on remaining DC's.

Dc17: syslog directory under /var/log/ is a soft link. On all other, syslog under /var/log is a mount point. This needs to be corrected, since the second syslog server can't be accessed (10.170.154.180) at this time. Did not fix the soft link issue on DC17.

Version1 of compress log script (compressLogV1.bash):

There exists a condition while compressing, if there exists an earlier compressed file.gz with the same name being compressed, the script would stall/fail waiting for input to overwrite the existing compressed file. This condition happened in DC4 and DC15.

As a work around checking for the existence of the compressed file with the same name and if it exists rotating (mv) the existing compressed file with current dateTime stamp so the next run of compression would not encounter this issue.

Version2 of compress log script (compressLogV2.bash):
The change was required to exclude ".snapshot" from the find, since DC16 has snapshot enabled for /var/log/syslog/ partition. This is the only DC thus far where snapshot is enabled. Until this is disabled, the fix is to exclude .snapshot directory from the find.

Version1 of purge log script (Purge90DasyLogV1.bash):

Cron Issue on SysLog Servers while Purging and Compressing logs

The change was required to exclude ".snapshot" from the find, since DC16 has snapshot enabled for /var/log/syslog/ partition. This is the only DC thus far where snapshot is enabled. Until this is disabled, the fix is to exclude .snapshot directory from the find.

Location of the Scripts: on each syslog (primary) server a directory called "syslog" is created under /roots to host the scripts.

Eg.

```
pc16splklog01:~/syslog # ls /root/syslog/*.bash
/root/syslog/compressLogV2.bash  /root/syslog/purge90daysLogV1.bash
pc16splklog01:~/syslog #
```

Location of the PurgeLog:

Both the scripts will keep a log of what was purged and what was compressed under purgelog directory. These scripts will create a directory under /var/log/syslog/, if it doesn't exist. Some stats are also appended to the log at the end of each run.

Below is the example of the stats for purgelog and compresslog.

```
pc16splklog01:/var/log/syslog/purgelog # ls
CompressLog-20180611  CompressLog-20180612  p90log-20180611  p90log-20180612

pc16splklog01:/var/log/syslog/purgelog # tail -4 p90log-20180611
Syslog Before Clean Up : Filesystem                                  Size  Used Avail Use% Mounted on
[2a00:0da9:0005:14b3:00e6:0000:0008:0007]:/SFA/vol_syslog_01 1000G  117G  884G  12% /var/log/
syslog
Syslog After Clean Up : Filesystem                                   Size  Used Avail Use% Mounted on
[2a00:0da9:0005:14b3:00e6:0000:0008:0007]:/SFA/vol_syslog_01 1000G   44G  957G   5% /var/log/
syslog

pc16splklog01:/var/log/syslog/purgelog # tail -4 CompressLog-20180612
Syslog Before Compression : Filesystem                               Size  Used Avail Use% Mounted on
[2a00:0da9:0005:14b3:00e6:0000:0008:0007]:/SFA/vol_syslog_01 1000G   44G  957G   5% /var/log/
syslog
Syslog After Compression : Filesystem                                Size  Used Avail Use% Mounted on
[2a00:0da9:0005:14b3:00e6:0000:0008:0007]:/SFA/vol_syslog_01 1000G  5.7G  995G   1% /var/log/
syslog
pc16splklog01:/var/log/syslog/purgelog #
```

Known caveat:

The new script will not check if the /var/log/syslog is 100% full, before writing the log to the partition. (Fix in future version).

Manjesh, Platform OSS team

Cron Issue on SysLog Servers while Purging and Compressing logs

If the syslog partition is not NFS mounted the script will exit (Happens only on DC17 syslog Server).

If the syslog partition has “snapshot” enabled, the script will parse thru that partition for both purge and compress scripts (Happens on DC16, Fix in future version [or] disable snapshot). (Fixed Jun12,2018)

<u>Pending work</u>:

1. Team to review the script (latest ones located on pc16splklog01) and validate for any runtime issues.
2. Remove the debug echo statements from the scripts
3. Make available the same version scripts at all the DC’s and then add the scripts as cron, keeping the same schedule as existing cleanup and compress.
4. Make available the scripts on the **<u>second syslogServer</u>** in every DC and update the cron, but hashed out, to prevent it from running twice.

Manjesh, Platform OSS team