openQA Infrastructure - action #53471

machine aarch64.o.o often unresponsive and needs power-cycle

2019-06-23 07:51 - okurz

Status:	Resolved		**Start date:**	2019-06-23
Priority:	Low		**Due date:**	
Assignee:	okurz		**% Done:**	0%
Category:			**Estimated time:**	0.00 hour
Target version:	Ready			

Description

Observation

aarch64.o.o had been unresponsive multiple times in the past weeksmonths, maybe after the upgrade to Leap 15.1 already since Leap 15.0 times?

I encountered this again 2019-06-23 09:10 CEST and gave the machine a power cycle with ipmitool -I lanplus -H openqa-aarch64-ipmi.suse.de -U ADMIN -P XXX power cycle as a remedy.

Checking /var/log/messages I can only find:

```
2019-06-22T03:30:00.073653+02:00 openqa-aarch64 worker[3005]: [info] quit due to signal TERM
2019-06-23T09:16:28.561376+02:00 openqa-aarch64 systemd[1]: systemd 234 running in system mode. (+
PAM -AUDIT +SELINUX -IMA +APPARMOR -SMACK +SYSVINIT +UTMP +LIBCRYPTSETUP +GCRYPT -GNUTLS +ACL +XZ
+LZ4 +SECCOMP +BLKID -ELFUTIL
```

so it looks like an openQA worker is terminated at 03:30, probably due to daily reboot command and the machine never comes up it seems?

Checking further log files with zgrep -B 1 'systemd .* running in system mode' /var/log/messages* | sed 's/systemd.*running in system mode.*$/BOOTUP/' | less I can find:

```
/var/log/messages-20190203.xz:2019-01-27T03:30:00.117876+01:00 openqa-aarch64 worker[2436]: [info]
 quit due to signal TERM
/var/log/messages-20190203.xz:2019-01-28T10:46:49.969565+01:00 openqa-aarch64 BOOTUP
```

which seems like the first occurence when the machine shut down but did not come up automatically. I assume someone on the next day power cycled it manually.

Taking a look in /var/log/zypp/history I do not find any suspicious packages like kernel update that could explain the system not coming up:

```
2019-01-27 00:42:46|command|root@openqa-aarch64|'zypper' '-R' '/tmp/tmp.pFe9WfeYHw' 'up' '-y' '--a
uto-agree-with-product-licenses'|
2019-01-27 00:42:47|install|openQA-common|4.6.1548532175.d411287e-lp150.1115.1|noarch||devel_openQ
A|7a31821029e964a30afb96c283d392f664f6d836d571eca98c82296caa91f756|
2019-01-27 00:42:47|install|openQA-client|4.6.1548532175.d411287e-lp150.1115.1|noarch||devel_openQ
A|fe261bc53a6b3ef0deb7ea1e24848a63dcdef07691601f716903ee3e3f7da893|
# 2019-01-27 00:42:47 openQA-worker-4.6.1548532175.d411287e-lp150.1115.1.noarch.rpm installed ok
# Additional rpm output:
# Running in chroot, ignoring request.
# Running in chroot, ignoring request.
#
2019-01-27 00:42:47|install|openQA-worker|4.6.1548532175.d411287e-lp150.1115.1|noarch||devel_openQ
A|7b6ad19005249fbc679f6f23d8940806942660380840af39f421cb14e2956cf3|
```

Problem

I suspect a problem on bootup causing – at best – a kernel crash which we could catch, installing kdump …

History

#1 - 2019-06-30 07:48 - okurz

2019-06-30 another incident, power cycled. I wonder about what would be a good option to monitor and apply the workaround of power cycling automatically. No record in /var/crash

#2 - 2019-07-02 06:27 - okurz

2019-07-02 again, power cycled. To debug if I can see anything while the system crashes or fails to come up (or maybe fails to shutdown properly?) I connected from lord.arch with SOL, currently running in screen, shell nr. 9, piping output into local/aarch64.o.o_sol_output_2019-07-02.log . Let's see if I can catch something the next crash.

#3 - 2019-07-02 07:05 - okurz

I plan to implement a [gitlab CI custom schedule](#) within https://gitlab.suse.de/openqa/monitoring , first trying out in https://gitlab.suse.de/okurz/monitoring , check for the system being up every day after the automatic upgrade and reboot should take care, e.g. at 04:20 CET/CEST in the morning or something.

#4 - 2019-07-02 13:17 - nicksinger

Have you been able to fully verify that the systems crash on their next bootup or does it happen earlier? With OSD's arm worker we face a similar issue described in here: https://progress.opensuse.org/issues/41882
We can link it to this one but only if it happens at runtime and not on reboot IMHO.

#5 - 2019-07-05 07:32 - okurz

nicksinger wrote:

> Have you been able to fully verify that the systems crash on their next bootup or does it happen earlier?

Earlier than "next bootup"? From the system logs as referenced in the description the only think I could find so far is:

```
/var/log/messages-20190203.xz:2019-01-27T03:30:00.117876+01:00 openqa-aarch64 worker[2436]: [info] quit due to
 signal TERM
/var/log/messages-20190203.xz:2019-01-28T10:46:49.969565+01:00 openqa-aarch64 BOOTUP
```

so anything that happens between the system receiving the signal to shutdown to the time when a new log appears after boot could be the problem, e.g. system fails to shutdown, fails in firmware, early kernel, systemd, etc. . Since [#53471#note-2](#) I have not seen a problem and the SOL session on lord.arch is still on. It has survived three reboots so far.

Side-note: In the SOL session I can see the following error messages immediately after the message Welcome to GRUB!:

```
error: serial port `com0' isn't found.
error: terminal `serial' isn't found.
error: terminal `serial' isn't found.
```

I guess that can be reported as grub bug if not already reported -> reported as https://bugzilla.opensuse.org/show_bug.cgi?id=1140464

#6 - 2019-07-06 05:07 - okurz

machine again stuck. As I am still in a SOL session I can see the last entries:

```
[  OK  ] Stopped target Local File Systems (Pre).
[  OK  ] Stopped Remount Root and Kernel File Systems.
         Stopping Monitoring of LVM2 mirrors…ng dmeventd or progress polling...
[  OK  ] Stopped Create Static Device Nodes in /dev.
[  OK  ] Reached target Shutdown.
```

so the machine is actually stuck in shutdown and fails to reboot from there. Sending [magic sysrq over IPMI SOL](#) has no effect though. Should I follow https://freedesktop.org/wiki/Software/systemd/Debugging/#index2h1 to debug shutdown problems?

Reported as https://bugzilla.opensuse.org/show_bug.cgi?id=1140572

Implemented a workaround in cron of user okurz on lord.arch to power cycle the machine if not responsive at 04:00 CET/CEST after daily update cycle.

#7 - 2019-07-10 15:11 - okurz

- Due date set to 2019-07-17

ok, after crosschecking with coolo in [#testing (RC)](#) I provided access to fbui as described in https://bugzilla.opensuse.org/show_bug.cgi?id=1140572#c9 . Let's give him a week maximum of debugging and then ask back. After this debugging approach is successful we can do the same for https://bugzilla.opensuse.org/show_bug.cgi?id=1138545

#8 - 2019-07-16 07:07 - okurz

- Due date changed from 2019-07-17 to 2019-07-24

fbui is currently debugging, let's give him some days

#9 - 2019-07-29 09:36 - okurz

- Due date deleted (2019-07-24)

- Status changed from Feedback to Blocked

- Priority changed from High to Low

I commented on [boo#1140572](#) that I enabled the automatic recovery script again after guillaume_g asked after the machine was unresponsive today again. Let's see if there is any solution in the bug, so "Blocked" here.

EDIT: 2019-10-24: I had systemd shutdown debugging running and collected some logs now, updated https://bugzilla.opensuse.org/show_bug.cgi?id=1140572#c17

EDIT: 2019-11-14: The problem looks related to NFS unmounting during shutdown. Asked in https://bugzilla.opensuse.org/show_bug.cgi?id=1140572#c23 what are plans as there seems to be an update for plymouth which might be related but it was done for SLE and not (yet) for Leap.

#10 - 2020-07-29 07:09 - okurz

- Target version set to Ready

#11 - 2020-08-05 08:33 - okurz

- Status changed from Blocked to In Progress

no progress on bugs but we are pretty stable with our current approach including the recovery. I should move the cron entries to a more common, non-personal place though. I created https://gitlab.suse.de/openqa/monitor-o3/ with a very simple .gitlab-ci.yml and https://build.opensuse.org/package/show/home:okurz:container/ipmitool-ping-nc-ssh and will try to get that done.

#12 - 2020-08-05 08:38 - okurz

- Status changed from In Progress to Resolved

Created pipeline "Automatic daily monitoring and recovery check of o3 infrastructure" in https://gitlab.suse.de/openqa/monitor-o3/-/pipeline_schedules/58/edit which triggers 0400Z daily. https://gitlab.suse.de/openqa/monitor-o3/-/pipelines/73347 is the last successful manual invocation.

Removed the according crontab entries from lord.arch:

```
26 # workaround for aarch64.o.o often not reachable, send power cycle if daily
27 # update reboot (around 03:30) does not bring the machine up again
28 # https://progress.opensuse.org/issues/53471
29 # okurz: 2019-07-16 disabled for now to allow fbui debugging
30 0 4 * * * ssh o3 'ping -q -c 1 aarch64 >/dev/null' || ipmitool -I lanplus -H openqa-aarch64-ipmi.suse.de -U
 ADMIN -P ADMIN power cycle
31 0 4 * * * ssh o3 'ping -q -c 1 openqaworker1 >/dev/null' || ipmitool -I lanplus -H openqaworker1-ipmi.suse.
de -U admin -P admin power cycle
32 0 4 * * * ssh o3 'ping -q -c 1 openqaworker4 >/dev/null' || ipmitool -I lanplus -H openqaworker4-ipmi.suse.
de -U admin -P admin power cycle
```