

## openQA Project - action #91764

### openqaworker13 has active jobs but not doing anything useful for hours

2021-04-26 13:45 - okurz

<b>Status:</b>	Resolved	<b>Start date:</b>	2021-04-26
<b>Priority:</b>	High	<b>Due date:</b>	2021-05-25
<b>Assignee:</b>	mkittler	<b>% Done:</b>	0%
<b>Category:</b>	Concrete Bugs	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	Ready		
<b>Difficulty:</b>			
<b>Description</b>			
<b>Observation</b>			
From <a href="https://chat.suse.de/channel/testing?msg=Rmq9EBMSJC7PyN8tt">https://chat.suse.de/channel/testing?msg=Rmq9EBMSJC7PyN8tt</a>			
Martin Doucha @mdoucha 15:39 Oliver Kurz I have a handful of jobs "running" on openqaworker13 for 11 hours already, is the whole machine down or something?			
okurz Oliver Kurz @okurz 15:42 <a href="https://stats.openqa-monitor.qa.suse.de/d/WDopenqaworker13/worker-dashboard-openqaworker13?orgId=1&amp;refresh=1m&amp;from=now-24h&amp;to=now">https://stats.openqa-monitor.qa.suse.de/d/WDopenqaworker13/worker-dashboard-openqaworker13?orgId=1&amp;refresh=1m&amp;from=now-24h&amp;to=now</a> looks more like it's bored  <a href="https://openqa.suse.de/admin/workers">https://openqa.suse.de/admin/workers</a> shows instances 1-16 "working", e.g. instance 1 reported working on <a href="https://openqa.suse.de/tests/5895608">https://openqa.suse.de/tests/5895608</a> which was cancelled some hours ago.			

## History

### #1 - 2021-04-26 13:47 - okurz

- Description updated

As we have persistent logs I will try to recover the machine now with a forced reboot but the issue should still be investigated.

### #2 - 2021-05-04 11:01 - cdywan

okurz wrote:

As we have persistent logs I will try to recover the machine now with a forced reboot but the issue should still be investigated.

So did you do something here? What's the state of it?

### #3 - 2021-05-04 13:13 - mkittler

- Assignee set to mkittler

I'll try to investigate the issue. Note that it would have increased my success *not* to reboot the machine immediately so I could observe the stuck processes. So I'll just dig into the logs then.

### #4 - 2021-05-04 13:39 - mkittler

- Status changed from New to Feedback

The logs don't go far enough to see what happened before the workers got stuck but it looks like the jobs were stuck during the setup phase waiting for the cache service, e.g. journalctl -u openqa-worker-auto-restart@2:

```
-- Logs begin at Sun 2020-12-20 03:30:00 CET, end at Tue 2021-05-04 15:17:06 CEST. --
```

```
Apr 26 10:59:06 openqaworker13 worker[32603]: [debug] [pid:32603] Updating status so job 5895608 is not consid
```

```

ered dead.
Apr 26 10:59:06 openqaworker13 worker[32603]: [debug] [pid:32603] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5895608/status
Apr 26 10:59:10 openqaworker13 worker[32603]: [debug] [pid:32603] Updating status so job 5895608 is not considered dead.
Apr 26 10:59:10 openqaworker13 worker[32603]: [debug] [pid:32603] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5895608/status
Apr 26 10:59:15 openqaworker13 worker[32603]: [debug] [pid:32603] Updating status so job 5895608 is not considered dead.
Apr 26 10:59:15 openqaworker13 worker[32603]: [debug] [pid:32603] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5895608/status
Apr 26 10:59:20 openqaworker13 worker[32603]: [debug] [pid:32603] Updating status so job 5895608 is not considered dead.
Apr 26 10:59:20 openqaworker13 worker[32603]: [debug] [pid:32603] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5895608/status
...
Apr 26 15:47:23 openqaworker13 worker[31821]: [debug] [pid:31821] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5894983/status
Apr 26 15:47:28 openqaworker13 worker[31821]: [debug] [pid:31821] Updating status so job 5894983 is not considered dead.
Apr 26 15:47:28 openqaworker13 worker[31821]: [debug] [pid:31821] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5894983/status
Apr 26 15:47:31 openqaworker13 worker[31821]: [info] [pid:31821] Received signal TERM
Apr 26 15:47:31 openqaworker13 worker[31821]: [debug] [pid:31821] Stopping job 5894983 from openqa.suse.de: 05894983-sle-15-Server-DVD-Updates-x86_64-Build20210426-1-qam-xfstests_bttrfs-generic-301-400@64bit - reason: quit
Apr 26 15:47:31 openqaworker13 worker[31821]: [debug] [pid:31821] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5894983/status
Apr 26 15:47:31 openqaworker13 systemd[1]: Stopping openQA Worker #2...
Apr 26 15:47:31 openqaworker13 worker[31821]: [info] [pid:31821] +++ worker notes +++
Apr 26 15:47:31 openqaworker13 worker[31821]: [info] [pid:31821] End time: 2021-04-26 13:47:31
Apr 26 15:47:31 openqaworker13 worker[31821]: [info] [pid:31821] Result: quit
Apr 26 15:47:31 openqaworker13 worker[31821]: [debug] [pid:31821] Job 5894983 stopped as incomplete
Apr 26 15:47:31 openqaworker13 worker[31821]: [debug] [pid:31821] Duplicating job 5894983
Apr 26 15:47:31 openqaworker13 worker[31821]: [debug] [pid:31821] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5894983/duplicate?dup_type_auto=1
Apr 26 15:47:31 openqaworker13 worker[31821]: [warn] [pid:31821] Unable to upload results of the job because no command server URL or worker ID have been set.
Apr 26 15:47:31 openqaworker13 worker[31821]: [debug] [pid:31821] Upload concluded (no current module)
Apr 26 15:47:31 openqaworker13 worker[31821]: [debug] [pid:31821] REST-API call: POST http://openqa.suse.de/api/v1/jobs/5894983/set_done?reason=quit%3A+worker+has+been+stopped+or+restarted&result=incomplete&worker_id=1113
Apr 26 15:47:32 openqaworker13 worker[31821]: [debug] [pid:31821] Job 5894983 from openqa.suse.de finished - reason: quit
Apr 26 15:47:32 openqaworker13 worker[31821]: [debug] [pid:31821] Informing openqa.suse.de that we are going offline

```

The logs for the other worker slots looks the same.

reported working on  
<https://openqa.suse.de/tests/5895608>  
which was cancelled some hours ago.

I assume the jobs were cancelled within the web UI because a user though cancelling might help to get rid of the stuck job. I suppose cancelling a job during the setup is not possible (except by sending a signal to the worker process). Now the jobs are incomplete because [okurz](#) restarted the workers.

The logs of the cacheservice(-minion) don't show anything interesting:

```

-- Logs begin at Sun 2020-12-20 03:30:00 CET, end at Tue 2021-05-04 15:25:03 CEST. --
Apr 26 11:37:47 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Stopping worker 12400 gracefully (80 seconds)
Apr 26 11:37:47 openqaworker13 openqa-workercache-daemon[9610]: [26005] [i] Worker 26005 started
Apr 26 11:37:51 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 12400 stopped
Apr 26 11:45:46 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Stopping worker 16696 gracefully (80 seconds)
Apr 26 11:45:46 openqaworker13 openqa-workercache-daemon[9610]: [26597] [i] Worker 26597 started
Apr 26 11:45:50 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 16696 stopped
Apr 26 12:23:07 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Stopping worker 19064 gracefully (80 seconds)
Apr 26 12:23:07 openqaworker13 openqa-workercache-daemon[9610]: [28296] [i] Worker 28296 started
Apr 26 12:23:09 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 19064 stopped
Apr 26 13:26:42 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Stopping worker 22869 gracefully (80 seconds)

```

```
Apr 26 13:26:42 openqaworker13 openqa-workercache-daemon[9610]: [32604] [i] Worker 32604 started
Apr 26 13:26:42 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 22869 stopped
Apr 26 14:40:58 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Stopping worker 26597 gracefully (8
0 seconds)
Apr 26 14:40:58 openqaworker13 openqa-workercache-daemon[9610]: [5283] [i] Worker 5283 started
Apr 26 14:41:02 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 26597 stopped
Apr 26 14:43:04 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Stopping worker 26005 gracefully (8
0 seconds)
Apr 26 14:43:04 openqaworker13 openqa-workercache-daemon[9610]: [5325] [i] Worker 5325 started
Apr 26 14:43:05 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 26005 stopped
Apr 26 15:14:19 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 28296 stopped
Apr 26 15:14:19 openqaworker13 openqa-workercache-daemon[9610]: [7166] [i] Worker 7166 started
Apr 26 15:31:28 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Stopping worker 32604 gracefully (8
0 seconds)
Apr 26 15:31:28 openqaworker13 openqa-workercache-daemon[9610]: [8121] [i] Worker 8121 started
Apr 26 15:31:29 openqaworker13 openqa-workercache-daemon[9610]: [9610] [i] Worker 32604 stopped
Apr 26 15:47:31 openqaworker13 systemd[1]: Stopping OpenQA Worker Cache Service Minion...
Apr 26 15:47:31 openqaworker13 openqa-worker-cacheservice-minion[9612]: rsync error: received SIGINT, SIGTERM,
or SIGHUP (code 20) at rsync.c(642) [generator=3.1.3]
Apr 26 15:47:31 openqaworker13 openqa-worker-cacheservice-minion[9612]: rsync error: received SIGUSR1 (code 19
) at main.c(1455) [receiver=3.1.3]
Apr 26 15:47:34 openqaworker13 openqa-worker-cacheservice-minion[9612]: [9612] [i] Worker 9612 stopped
Apr 26 15:47:34 openqaworker13 systemd[1]: openqa-worker-cacheservice-minion.service: Main process exited, cod
e=exited, status=16/n/a
Apr 26 15:47:34 openqaworker13 systemd[1]: Stopped OpenQA Worker Cache Service Minion.
Apr 26 15:47:34 openqaworker13 systemd[1]: openqa-worker-cacheservice-minion.service: Unit entered failed stat
e.
Apr 26 15:47:34 openqaworker13 systemd[1]: openqa-worker-cacheservice-minion.service: Failed with result 'exit
-code'.
Apr 26 15:47:34 openqaworker13 systemd[1]: Stopping OpenQA Worker Cache Service...
Apr 26 15:47:34 openqaworker13 openqa-workercache-daemon[9610]: [9610] [w] Stopping worker 8121 immediately
Apr 26 15:47:34 openqaworker13 openqa-workercache-daemon[9610]: [9610] [w] Stopping worker 7166 immediately
Apr 26 15:47:34 openqaworker13 openqa-workercache-daemon[9610]: [9610] [w] Stopping worker 5325 immediately
```

The full journal doesn't contain anything interesting as well.

---

I would figure that the cacheservice was somehow stuck leading to all worker slots getting stuck at some point. However, the cacheservice was not completely stuck. It just didn't process the jobs but apparently still responded to HTTP requests and workers didn't complain that the cacheservice is entirely unavailable (and hence just got stuck). Maybe a lock was stuck or all Minion workers were stuck? The Minion database is cleared when rebooting so there are unfortunately no traces left to investigate this further.

I suppose I had to look on the worker at the time we found out about it but now it is not possible to tell what was going on.

#### #5 - 2021-05-05 17:46 - okurz

I suppose I had to look on the worker at the time we found out about it but now it is not possible to tell what was going on.

any thoughts on checks we could build into our source code in the meantime? Maybe something like timeouts for certain tasks and if a timeout is hit and a task has not finished doing something useful we raise that as an error?

#### #6 - 2021-05-06 15:23 - mkittler

A timeout would make sense (similar or in-line with the existing timeout via MAX\_JOB\_TIME). It would also be useful if the job was cancellable by the web UI in this case.

#### #7 - 2021-05-10 16:40 - mkittler

- Status changed from Feedback to In Progress

PR: <https://github.com/os-autoinst/openQA/pull/3894>

#### #8 - 2021-05-11 04:10 - openqa\_review

- Due date set to 2021-05-25

Setting due date based on mean cycle time of SUSE QE Tools

#### #9 - 2021-05-17 08:28 - mkittler

- Status changed from In Progress to Feedback

I've been merging the PR. So now there's MAX\_SETUP\_TIME (by default one hour) and manual cancelling works as well while waiting for the cache service.

**#10 - 2021-05-18 10:46 - okurz**

- *Status changed from Feedback to Resolved*

merged and deployed, MAX\_SETUP\_TIME should be effective. There are currently no alerts and in case of regressions we expect alerts, e.g. job age either because jobs are stuck on workers or because jobs are stuck in scheduling because we exhaust all free workers :)