

## openQA Project - action #16380

### workers die like flies

2017-02-01 06:13 - coolo

<b>Status:</b>	Rejected	<b>Start date:</b>	2017-02-01
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	oholecek	<b>% Done:</b>	80%
<b>Category:</b>	Concrete Bugs	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	Milestone 8		
<b>Difficulty:</b>			
<b>Description</b>			
Most workers on openqa.opensuse.org are dead:			
Jan 31 12:46:47 openqaworker4 worker[13393]: can't open /var/lib/openqa/pool/15/testresults/test_order.json: No such file or directory at /usr/share/openqa/script/./lib/OpenQA/Worker/Jobs.pm line 735.			
Jan 31 12:46:52 openqaworker4 worker[13393]: [DEBUG] Either there is no job running or we were asked to stop: (1 Reason: no tests scheduled)			
Jan 31 12:46:52 openqaworker4 worker[13393]: [INFO] cleaning up 00343867-opensuse-42.3-DVD-x86_64-Build0055-rescue_system@uefi			
Jan 31 12:46:52 openqaworker4 worker[13393]: [INFO] got job 343869: 00343869-opensuse-42.3-DVD-x86_64-Build0055-textmode-image@64bit			
Jan 31 12:46:52 openqaworker4 worker[13393]: [INFO] 6853: WORKING 343869			
Jan 31 12:49:35 openqaworker4 worker[13393]: [ERROR] 404 response: Not Found (remaining tries: 0)			
Jan 31 12:49:35 openqaworker4 worker[13393]: [ERROR] Job aborted because web UI doesn't accept updates anymore (likely considers this job dead)			
Jan 31 12:49:36 openqaworker4 worker[13393]: Mojo::Reactor::Poll: Timer failed: No worker id or webui host set! at /usr/share/openqa/script/./lib/OpenQA/Worker/Common.pm line 181.			
Jan 31 12:49:45 openqaworker4 worker[13393]: WebUI Mojo::IOLoop=HASH(0x37719a0) is unknown! - Should not happen but happened, exiting! at /usr/share/openqa/script/./lib/OpenQA/Worker/Common.pm line 404.			
Jan 31 12:49:45 openqaworker4 worker[13393]: [INFO] registering worker with openQA Mojo::IOLoop=HASH(0x37719a0)...			

### History

#### #1 - 2017-02-01 10:31 - oholecek

This happened to me when job was set done before all status updates were finished, we need to review `stop_job` ordering. Second part is that the response on unsuccessful api call is too harsh I think. 404 should not end as an api failure and shutting down the worker IMO.

#### #2 - 2017-02-01 11:37 - coolo

No, but it should stop the job. Dead jobs are marked this way.

#### #3 - 2017-02-01 15:28 - oholecek

- Status changed from New to In Progress

- Assignee set to oholecek

There is obvious bug in scheduling worker registration without `$host` specified and `stop_job` clearing the `$current_host` in the meantime. Testing fix for it now. But this fixes the symptom - dying worker, not the actual cause, which will still trigger marking the job as incomplete.

#### #4 - 2017-03-01 09:24 - okurz

- Target version changed from Milestone 5 to Milestone 6

#### #5 - 2017-03-14 19:30 - AdamWill

The symptom is pretty bad and needs fixing regardless (this is also affecting both Fedora openQA deployments and killing our workers).

But yes, we do need to also figure out *why* workers are getting 404s in the first place. Here's the logs from one case on our staging instance.

Worker log:

Mar 14 17:18:12 qa09.qa.fedoraproject.org worker[5943]: [INFO] 19431: WORKING 79745

Mar 14 17:23:01 qa09.qa.fedoraproject.org worker[5943]: [ERROR] 404 response: Not Found (remaining tries: 0)

Mar 14 17:23:01 qa09.qa.fedoraproject.org worker[5943]: [ERROR] Job aborted because web UI doesn't accept updates anymore (likely considers

this job dead)

Mar 14 17:23:01 qa09.qa.fedoraproject.org worker[5943]: Mojo::Reactor::Poll: Timer failed: No worker id or webui host set! at /usr/share/openqa/script/./lib/OpenQA/Worker/Common.pm line 181.

Mar 14 17:23:03 qa09.qa.fedoraproject.org worker[5943]: WebUI Mojo::IOLoop=HASH(0x3558158) is unknown! - Should not happen but happened, exiting! at /usr/share/openqa/script/./lib/OpenQA/Worker/Common.pm line 404.

Mar 14 17:23:03 qa09.qa.fedoraproject.org worker[5943]: [INFO] registering worker with openQA Mojo::IOLoop=HASH(0x3558158)...

Server Apache log, grepped for 79745:

```
[root@openqa-stg01 adamwill][PROD]# grep 79745 /var/log/httpd/access_log
10.5.124.239 - - [14/Mar/2017:17:18:22 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:18:32 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:18:42 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:18:52 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:02 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:12 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:22 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 261 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:32 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 226 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:42 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 191 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:43 +0000] "POST /api/v1/jobs/79745/artefact HTTP/1.1" 200 2 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:44 +0000] "POST /api/v1/jobs/79745/artefact HTTP/1.1" 200 2 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:19:52 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:20:02 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:20:12 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:20:22 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:20:32 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:20:42 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:20:52 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:21:02 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:21:12 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:21:22 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:21:32 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:21:42 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:21:52 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 200 52 "-" "Mojolicious (Perl)"
10.5.124.239 - - [14/Mar/2017:17:22:52 +0000] "POST /api/v1/jobs/79745/status HTTP/1.1" 404 49 "-" "Mojolicious (Perl)"
```

Server journal, grepped for 79745:

Mar 14 17:22:44 openqa-stg01.qa.fedoraproject.org openqa-websockets[20266]: [Tue Mar 14 17:22:44 2017] [websockets:debug] job considered dead: 79745 worker 24 not seen. In state running

Mar 14 17:22:45 openqa-stg01.qa.fedoraproject.org openqa-websockets[20266]: [Tue Mar 14 17:22:45 2017] [websockets:warn] dead job 79745 aborted and duplicated 79978

Mar 14 17:22:52 openqa-stg01.qa.fedoraproject.org openqa[20267]: [Tue Mar 14 17:22:52 2017] [13376:info] Got status update for job with no worker assigned (maybe running job already considered dead): 79745

So basically it seems like the worker is checking in every ten seconds and everything is hunky dory, then the worker neglects to check in for a minute (between 17:21:52 and 17:22:52). During that minute, websockets decides the job is dead - at 17:22:44 - and aborts it. Then when the worker *does* try and check in at 17:22:52, the server returns a 404 and logs it. And that's the 404 that triggers the worker death due to the issue with the `register_worker` call.

#### #6 - 2017-03-14 20:12 - AdamWill

Wow, OK, so something wacky clearly happened on the worker host in question for this case (and one other) around the time in question...there's some messy complaints from the kernel. Not sure exactly what was going on, but logs attached for interest. Dunno if this is what's happening to you guys, or if it's something different.

#### #7 - 2017-03-14 20:13 - AdamWill

- *File wtf.txt added*

#### #8 - 2017-03-20 13:13 - oholecek

- *% Done changed from 0 to 80*

[PB](#) merged, the tests are still missing.

#### #9 - 2017-03-23 10:08 - okurz

- *Priority changed from High to Normal*

I guess this is less critical now. reducing prio.

#### #10 - 2017-04-20 09:12 - okurz

- Target version changed from Milestone 6 to Milestone 7

bump

**#11 - 2017-05-23 06:48 - okurz**

- Target version changed from Milestone 7 to Milestone 8

Who has information if this still applies? Are there any workaround cronjob/monitoring jobs that restart workers because of this or are we done here?

**#12 - 2017-05-23 07:14 - coolo**

if you could stop messing with target versions outside of backlog planning meetings, we would have a way to go through this systemetically. But continue messing

**#13 - 2017-07-17 07:28 - coolo**

- Status changed from In Progress to Rejected

the github issue is there, but left WIP for months

## Files

---

wtf.txt	424 KB	2017-03-14	AdamWill
---------	--------	------------	----------