

openQA Project - action #59043

Fix unstable/flaky full-stack test, i.e. remove sleep, and ui tests

2019-11-04 16:25 - okurz

Status: Resolved	Start date: 2019-11-04
Priority: Low	Due date:
Assignee: okurz	% Done: 0%
Category: Concrete Bugs	Estimated time: 0.00 hour
Target version: Ready	
Difficulty:	

Description

Motivation

<https://github.com/os-autoinst/openQA/pull/2455> introduced a sleep 8 to stabilize a test. Having sleeps in code is better than unstable tests but even better would be to have reliable code that explicitly synchronizes on necessary conditions rather than blind sleeps. Also we have more flaky/unstable tests which we maintain in `.circleci/unstable_tests.txt`.

Acceptance criteria

- **AC1:** DONE: ~~sleep 8 of <https://github.com/os-autoinst/openQA/pull/2455> is removed~~ ~* **AC2:** `.circleci/unstable_tests.txt` is empty again ~* ## Suggestions
- Remove explicit and implicit sleep time, e.g. as in <https://github.com/os-autoinst/openQA/pull/2626> for the 2s selenium test method time
- Run tests from `.circleci/unstable_tests.txt` locally, reproduce problems and fix them
- Ensure tests also run stable within CI, not only locally

Related issues:

Related to openQA Project - action #59100: Race condition in os-autoinst/comm...	Rejected	2019-11-05
Related to openQA Project - action #62015: jobs incomplete without logs as so...	Resolved	2020-01-10
Related to openQA Project - action #54032: flaky test: t/full-stack.t "not ok...	Resolved	2019-07-09
Related to openQA Project - action #68836: t/api/14-plugin_obs_rsync_async.t ...	Workable	2020-07-10
Related to openQA Project - action #71554: unstable/flaky/sporadic t/full-sta...	Resolved	2020-09-19
Related to openQA Project - action #76900: unstable/flaky/sporadic t/full-sta...	Resolved	
Copied to openQA Project - action #72085: flaky/unstable t/05-scheduler-full...	Workable	
Copied to openQA Project - action #88496: openQA "t" tests time out (again) a...	Resolved	2019-11-04 2021-02-23

History

#1 - 2019-11-05 10:55 - okurz

- Related to action #59100: Race condition in os-autoinst/commands.pm added

#2 - 2019-11-06 10:52 - tinita

I'll add what I found out so far.

The failure can be reproduced locally by adding a sleep in isotovideo:

```
sleep 8;
($testprocess, $testfd) = autotest::start_process();
```

Debugging shows that the `set_pause_at_test` command is going through `send_json` before isotovideo starts to read:

```
# enter the main loop: process messages from autotest, command server and backend
while ($loop) {
```

If I add back the HTTP status call in `openQA/lib/OpenQA/Worker/Job.pm` `_upload_results`, then it is sent after isotovideo starts reading.

#3 - 2019-11-07 08:30 - coolo

- Priority changed from Normal to Low

I see other issues as more urgent

#4 - 2020-01-17 05:31 - okurz

- Related to action #62015: jobs incomplete without logs as some workers are rejected (was: Scheduler does not work) added

#5 - 2020-01-24 13:22 - okurz

- Subject changed from Fix full-stack test flakyness, i.e. remove sleep to Fix unstable/flaky full-stack test, i.e. remove sleep

#6 - 2020-01-24 13:22 - okurz

- Related to action #54032: flaky test: t/full-stack.t "not ok 44 - Cache tests", subtest: "not ok 17 - test 6 is passed" at /opt/testing_area/openqa/t/lib/OpenQA/Test/FullstackUtils.pm line 120. added

#7 - 2020-01-24 13:23 - okurz

- Status changed from New to Feedback
- Assignee set to okurz

Checking stability in <https://github.com/os-autoinst/openQA/pull/2496>

#8 - 2020-02-06 08:22 - cdywan

- Target version set to Current Sprint

#9 - 2020-02-24 07:34 - okurz

- Status changed from Feedback to Workable

First <https://github.com/os-autoinst/openQA/pull/2765>, then <https://github.com/os-autoinst/openQA/pull/2626#issuecomment-580173772>: it seems the retry and stability tests flag are not properly exported into the sub-scope of the retry script.

#10 - 2020-03-24 14:44 - okurz

<https://github.com/os-autoinst/openQA/pull/2811> is the fix for the retry script not handling "stability tests" properly. Now I can look into the test failures of <https://github.com/os-autoinst/openQA/pull/2626> again.

#11 - 2020-04-07 08:08 - okurz

It's hard to keep track on this. However, latest result from rebased code from yesterday show that all tests are very stable except for ui tests, e.g. <https://circleci.com/api/v1.1/project/github/os-autoinst/openQA/20110/output/108/0?file=true&allocation-id=5e8b4e234cffa121cb18c895-0-build%2FA7C7C4BF> showing

```
[17:16:00] t/ui/01-list.t ..... 19/? Bailout called. Further testing stopped: findElement: no such element: Unable to locate element: {"method":"css selector","selector":"#scheduled #job_99928"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:108
FAILED--Further testing stopped: findElement: no such element: Unable to locate element: {"method":"css selector","selector":"#scheduled #job_99928"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:108
Retrying...
[17:16:25] t/ui/01-list.t .....
[17:16:25] t/ui/01-list.t ..... 1/?
[17:16:25] t/ui/01-list.t ..... 12/?
[17:16:25] t/ui/01-list.t ..... 19/? Bailout called. Further testing stopped: findElement: no such element: Unable to locate element: {"method":"css selector","selector":"#scheduled #job_99928"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:108
FAILED--Further testing stopped: findElement: no such element: Unable to locate element: {"method":"css selector","selector":"#scheduled #job_99928"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:108
Retrying...
[17:16:43] t/ui/01-list.t .....
[17:16:43] t/ui/01-list.t ..... 1/?
[17:16:43] t/ui/01-list.t ..... 4/?
[17:16:43] t/ui/01-list.t ..... 14/?
[17:16:43] t/ui/01-list.t ..... 20/? Bailout called. Further testing stopped: findElement: no such element: Unable to locate element: {"method":"css selector","selector":"#scheduled #job_99928"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:108
FAILED--Further testing stopped: findElement: no such element: Unable to locate element: {"method":"css selector","selector":"#scheduled #job_99928"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:108
Retrying...
```

#12 - 2020-04-16 12:01 - okurz

- Status changed from Workable to In Progress

Latest PR <https://github.com/os-autoinst/openQA/pull/2945> , already merged. Continuing in <https://github.com/os-autoinst/openQA/pull/2626> with the goal to also remove all implicit selenium waits which helps us to identify the points where we need to (explicitly) wait and also reduce overall test time in case we do not actually need to wait.

#13 - 2020-04-17 13:55 - okurz

- Subject changed from *Fix unstable/flaky full-stack test, i.e. remove sleep* to *Fix unstable/flaky full-stack test, i.e. remove sleep, and ui tests*

<https://github.com/os-autoinst/openQA/pull/2949> for t/ui/10-tests_overview.t, next: <https://github.com/os-autoinst/openQA/pull/2952>

#14 - 2020-04-20 04:31 - okurz

After some problems in t/ui/13-admin.t I reverted one part with <https://github.com/os-autoinst/openQA/pull/2958> . Problem with that specific test module is also that locally I can never run it successfully but get different problems:

```
t/ui/13-admin.t .. 11/?
# Failed test 'group properties save button is disabled if name is left empty'
# at t/ui/13-admin.t line 437.
# got: undef
# expected: 'true'

# Failed test 'editor name input marked as invalid'
# at t/ui/13-admin.t line 439.
# got: 'form-control'
# expected: 'form-control is-invalid'

# Failed test 'group properties save button is disabled if name is blank'
# at t/ui/13-admin.t line 451.
# got: undef
# expected: 'true'

# Failed test 'editor name input marked as invalid'
# at t/ui/13-admin.t line 453.
# got: 'form-control'
# expected: 'form-control is-invalid'
# Looks like you failed 4 tests of 5.

# Failed test 'update group name with empty or blank'
# at t/ui/13-admin.t line 461.

# Failed test 'keep important results in days edited'
# at t/ui/13-admin.t line 483.
# got: '5000'
# expected: '500'
# Looks like you failed 1 test of 10.

# Failed test 'edit some properties'
# at t/ui/13-admin.t line 495.
# Looks like you failed 2 tests of 4.
```

#15 - 2020-04-23 15:52 - mkittler

It works for me locally (not only with my helpers, also with make test TESTS=t/ui/13-admin.t). The test might be disturbed by leftover files in your repository. You could easily check that by running the test on a clean checkout.

#16 - 2020-04-24 11:24 - okurz

<https://app.circleci.com/pipelines/github/os-autoinst/openQA/2501/workflows/3a8e5f9a-3e77-48fd-8d79-2dfc1a115860/jobs/23520>

shows:

```
[10:58:38] t/ui/01-list.t ..... 53/?
# Failed test 'only softfailed and failed jobs displayed'
# at t/ui/01-list.t line 342.
# Structures begin differing at:
# $got->[5] = Does not exist
# $expected->[5] = 'job_99944'
[10:58:38] t/ui/01-list.t ..... 59/?
# Failed test '99945 again hidden'
# at t/ui/01-list.t line 355.
# Structures begin differing at:
# $got->[8] = 'job_99945'
# $expected->[8] = 'job_80000'
```

#17 - 2020-04-26 19:32 - okurz

<https://github.com/os-autoinst/openQA/pull/2987> created for t/ui/01-list.t

Regarding the full stack test, I found out that there is a section of test code that is never executed in circleCI. That explains why tests always fail locally :(

<https://github.com/os-autoinst/openQA/pull/2997> for a fix of the failing test part.

What we should still try is to run all the tests on circleCI as well of course, see <https://github.com/os-autoinst/openQA/pull/2998> for that.

#18 - 2020-05-02 20:07 - okurz

- Status changed from In Progress to Feedback

<https://github.com/os-autoinst/openQA/pull/2987> merged to mark t/ui/01-list.t as unstable. To be "done" here we need to stabilize the test again so that we can remove it from "unstable" as well again. <https://github.com/os-autoinst/openQA/pull/2997> was merged to fix the local runs of the full stack test and <https://github.com/os-autoinst/openQA/pull/2998> was merged to ensure all code within t/full-stack.t is also executed within CI and all environments, regardless of any settings. This should avoid further bad surprises. <https://github.com/os-autoinst/openQA/pull/3027> created to cover some minor enhancements together with the actual original goal of removing the "sleep 8" from t/full-stack.t again. Additionally now trying again how current stability looks like in <https://github.com/os-autoinst/openQA/pull/2496>

#19 - 2020-05-04 21:09 - okurz

Lately "scheduler" tests are getting more unstable. I locally ran for i in {001..100}; do echo "## Run \$i / 100" && env SCHEDULER_FULLSTACK=1 FULLSTACK=1 KEEP_DB=1 STABILITY_TEST=1 timeout 60 prove -l t/05-scheduler-full.t; done and found

```
## Run 022 / 100
t/05-scheduler-full.t .. Can't create listen socket: Address already in use at /usr/lib/perl5/vendor_perl/5.26.1/Mojo/IOLoop.pm line 124.
t/05-scheduler-full.t .. 1/?
#   Failed test 'jobs dispatched to different workers'
#   at t/05-scheduler-full.t line 100.
#       got: undef
#       expected: anything else

#   Failed test 'each of the two jobs allocated to one of the workers'
#   at t/05-scheduler-full.t line 101.
#       got: undef
#       expected: anything else
# [
#   {},
#   {}
# ]
# Looks like you failed 2 tests of 4.
t/05-scheduler-full.t .. 2/?
#   Failed test 'Scheduler worker job allocation'
#   at t/05-scheduler-full.t line 109.

#   Failed test 'one job allocated'
#   at t/05-scheduler-full.t line 137.
#       got: '0'
#       expected: '1'
# Looks like you failed 1 test of 3.
t/05-scheduler-full.t .. 3/?
#   Failed test 're-scheduling and incompletion of jobs when worker rejects jobs or goes offline'
#   at t/05-scheduler-full.t line 189.
Can't use an undefined value as a HASH reference at t/05-scheduler-full.t line 138.
# Tests were run but no plan was declared and done_testing() was not seen.
...
## Run 055 / 100
t/05-scheduler-full.t .. 2/?
#   Failed test 'scheduler does not consider broken worker for allocating job'
#   at t/05-scheduler-full.t line 128.
#       got: '1'
#       expected: '0'

#   Failed test 'one job allocated'
#   at t/05-scheduler-full.t line 163.
#       got: '0'
#       expected: '1'
# Looks like you failed 2 tests of 7.
t/05-scheduler-full.t .. 3/?
#   Failed test 're-scheduling and incompletion of jobs when worker rejects jobs or goes offline'
#   at t/05-scheduler-full.t line 189.
```

```

Can't use an undefined value as a HASH reference at t/05-scheduler-full.t line 164.
# Tests were run but no plan was declared and done_testing() was not seen.
...
## Run 066 / 100
t/05-scheduler-full.t .. Can't create listen socket: Address already in use at /usr/lib/perl5/vendor_perl/5.26
.1/Mojo/IOLoop.pm line 124.
t/05-scheduler-full.t .. 1/?
#   Failed test 'jobs dispatched to different workers'
#   at t/05-scheduler-full.t line 100.
#       got: undef
#   expected: anything else

#   Failed test 'each of the two jobs allocated to one of the workers'
#   at t/05-scheduler-full.t line 101.
#       got: undef
#   expected: anything else
# [
#   {},
#   {}
# ]
# Looks like you failed 2 tests of 4.
t/05-scheduler-full.t .. 2/?
#   Failed test 'Scheduler worker job allocation'
#   at t/05-scheduler-full.t line 109.

#   Failed test 'one job allocated'
#   at t/05-scheduler-full.t line 137.
#       got: '0'
#   expected: '1'
# Looks like you failed 1 test of 3.
t/05-scheduler-full.t .. 3/?
#   Failed test 're-scheduling and incompleteness of jobs when worker rejects jobs or goes offline'
#   at t/05-scheduler-full.t line 189.
Can't use an undefined value as a HASH reference at t/05-scheduler-full.t line 138.
# Tests were run but no plan was declared and done_testing() was not seen.
...
## Run 088 / 100
t/05-scheduler-full.t .. 5/? ## Run 089 / 100
t/05-scheduler-full.t .. Can't create listen socket: Address already in use at /usr/lib/perl5/vendor_perl/5.26
.1/Mojo/IOLoop.pm line 124.
t/05-scheduler-full.t .. 1/?
#   Failed test 'jobs dispatched to different workers'
#   at t/05-scheduler-full.t line 100.
#       got: undef
#   expected: anything else

#   Failed test 'each of the two jobs allocated to one of the workers'
#   at t/05-scheduler-full.t line 101.
#       got: undef
#   expected: anything else
# [
#   {},
#   {}
# ]
# Looks like you failed 2 tests of 4.
t/05-scheduler-full.t .. 2/?
#   Failed test 'Scheduler worker job allocation'
#   at t/05-scheduler-full.t line 109.

#   Failed test 'one job allocated'
#   at t/05-scheduler-full.t line 137.
#       got: '0'
#   expected: '1'
# Looks like you failed 1 test of 3.
t/05-scheduler-full.t .. 3/?
#   Failed test 're-scheduling and incompleteness of jobs when worker rejects jobs or goes offline'
#   at t/05-scheduler-full.t line 189.
Can't use an undefined value as a HASH reference at t/05-scheduler-full.t line 138.
# Tests were run but no plan was declared and done_testing() was not seen.
...
## Run 094 / 100
t/05-scheduler-full.t .. Can't create listen socket: Address already in use at /usr/lib/perl5/vendor_perl/5.26
.1/Mojo/IOLoop.pm line 124.
t/05-scheduler-full.t .. 1/?
#   Failed test 'jobs dispatched to different workers'

```

```

# at t/05-scheduler-full.t line 100.
# got: undef
# expected: anything else

# Failed test 'each of the two jobs allocated to one of the workers'
# at t/05-scheduler-full.t line 101.
# got: undef
# expected: anything else
# [
# {},
# {}
# ]
# Looks like you failed 2 tests of 4.
t/05-scheduler-full.t .. 2/?
# Failed test 'Scheduler worker job allocation'
# at t/05-scheduler-full.t line 109.

# Failed test 'one job allocated'
# at t/05-scheduler-full.t line 137.
# got: '0'
# expected: '1'
# Looks like you failed 1 test of 3.
t/05-scheduler-full.t .. 3/?
# Failed test 're-scheduling and incompletion of jobs when worker rejects jobs or goes offline'
# at t/05-scheduler-full.t line 189.
Can't use an undefined value as a HASH reference at t/05-scheduler-full.t line 138.
# Tests were run but no plan was declared and done_testing() was not seen.
...
## Run 098 / 100
t/05-scheduler-full.t .. 2/?
# Failed test 'scheduler does not consider broken worker for allocating job'
# at t/05-scheduler-full.t line 128.
# got: '1'
# expected: '0'
# Looks like you failed 1 test of 12.
t/05-scheduler-full.t .. 3/?
# Failed test 're-scheduling and incompletion of jobs when worker rejects jobs or goes offline'
# at t/05-scheduler-full.t line 189.
t/05-scheduler-full.t .. 5/? # Looks like you failed 1 test of 5.

```

so 4/100 times "Can't create listen socket: Address already in use at /usr/lib/perl5/vendor_perl/5.26.1/Mojo/IOLoop.pm line 124." and 2/100 times "Failed test 'scheduler does not consider broken worker for allocating job' at t/05-scheduler-full.t line 128."

EDIT: 2020-05-05: Discussed with mktittler and kraih. Mojo::IOLoop::Server::generate_port is a nice thing as it looks for a port that can actually be bound to but it does not reserve the port so there might be a race condition when by chance another service binds to the port between the call to generate_port and actually listening on it in a service. It might be possible to still use an IO::Socket::IP object and pass that object to any using service to keep it reserved. Upstream mojo uses this idiom now <https://github.com/mojolicious/mojo/blob/master/t/mojo/daemon.t#L261-L263>, it lets the server bind to a random port and then the server provides an api for retrieving the port, never releasing the port, means you have to bind before fork though. I think we should be able to adapt our code to not rely implicitly on a sequential port range anyway. This is also related to [#49883](#) as well.

#20 - 2020-05-05 07:26 - okurz

from <https://app.circleci.com/pipelines/github/os-autoinst/openQA/2688/workflows/40eafa8f-32f5-47cc-8cbc-ea2028be05cd/jobs/25303> on master:

```

[05:14:14] t/ui/26-jobs_restart.t ..... 7/? Bailout called. Further testing stopped: clickElement: element click intercepted: Element <a class="cancel" href="/api/v1/jobs/99986/cancel">...</a> is not clickable at point (359, 591). Other element would receive the click: <td class=" test">...</td> at /home/squama ta/project/t/ui/./lib/OpenQA/SeleniumTest.pm:109

```

```

FAILED--Further testing stopped: clickElement: element click intercepted: Element <a class="cancel" href="/api/v1/jobs/99986/cancel">...</a> is not clickable at point (359, 591). Other element would receive the click: <td class=" test">...</td> at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:109

```

#21 - 2020-05-06 19:07 - okurz

Trying to bring in the changes to make the "test pausing" more stable as well as remove the sleep 8. It seems as in my local environment, with kvm available, tests are more likely to "miss" the moment to stop the test so removing sleep 8 helps but on circle removing the sleep 8 causes other problems.

I managed to prepare a change to set "PAUSE_AT" at the first job to be scheduled. With proper job settings the other jobs work fine after the first one in the full stack test that we want to "pause" but the first job is waiting for 4 minutes before it is continued:

```

[2020-05-06T20:54:11.0098 CEST] [debug] [pid:20641] Running job 1 from http://localhost:42237: 00000001-tinycore-1-flavor-i386-Build1-core@coolone.
[2020-05-06T20:54:11.0100 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s

```

```
tatus
[2020-05-06T20:54:15.0476 CEST] [debug] [pid:20641] Starting livelog
[2020-05-06T20:54:15.0477 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s
tatus
[2020-05-06T20:54:15.0633 CEST] [debug] [pid:20641] Stopping livelog
[2020-05-06T20:54:16.0544 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s
tatus
[2020-05-06T20:54:26.0590 CEST] [debug] [pid:20641] Unable to read result-shutdown.json: No such file or direc
tory
[2020-05-06T20:54:26.0592 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s
tatus
[2020-05-06T20:54:26.0689 CEST] [debug] [pid:20695] Optimizing /tmp/PplAxYOqSM/t/full-stack.d/openqa/pool/1/te
stresults/boot-2.png
[2020-05-06T20:54:26.0765 CEST] [debug] [pid:20695] Uploading artefact boot-2.png as 68d7ac83452b72555f8a6b386
629f91f
[2020-05-06T20:54:26.0786 CEST] [debug] [pid:20695] Optimizing /tmp/PplAxYOqSM/t/full-stack.d/openqa/pool/1/te
stresults/.thumbs/boot-2.png
[2020-05-06T20:54:26.0791 CEST] [debug] [pid:20695] Uploading artefact boot-2.png as 68d7ac83452b72555f8a6b386
629f91f
[2020-05-06T20:54:26.0811 CEST] [debug] [pid:20695] Optimizing /tmp/PplAxYOqSM/t/full-stack.d/openqa/pool/1/te
stresults/boot-1.png
[2020-05-06T20:54:26.0902 CEST] [debug] [pid:20695] Uploading artefact boot-1.png as ffc79660bfb954ce25160084
5c86885
[2020-05-06T20:54:26.0919 CEST] [debug] [pid:20695] Optimizing /tmp/PplAxYOqSM/t/full-stack.d/openqa/pool/1/te
stresults/.thumbs/boot-1.png
[2020-05-06T20:54:26.0924 CEST] [debug] [pid:20695] Uploading artefact boot-1.png as ffc79660bfb954ce25160084
5c86885
[2020-05-06T20:54:36.0949 CEST] [debug] [pid:20641] Unable to read result-shutdown.json: No such file or direc
tory
...
[2020-05-06T20:59:18.0096 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s
tatus
[2020-05-06T20:59:28.0137 CEST] [debug] [pid:20641] Unable to read result-shutdown.json: No such file or direc
tory
[2020-05-06T20:59:28.0139 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s
tatus
[2020-05-06T20:59:30.0860 CEST] [debug] [pid:20641] Starting livelog
[2020-05-06T20:59:30.0861 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s
tatus
[2020-05-06T20:59:31.0910 CEST] [debug] [pid:20641] REST-API call: POST http://localhost:42237/api/v1/jobs/1/s
tatus
```

#22 - 2020-05-07 18:12 - okurz

<https://github.com/os-autoinst/openQA/pull/3039> merged to remove the "sleep 8", but

<https://app.circleci.com/pipelines/github/os-autoinst/openQA/2766/workflows/83d9f777-5f5c-4926-a148-2b04aa12f1bd/jobs/26094/steps> shows that run 8 of a stability test run for fullstack failed with

Rerun 8 of 20 ...

```
[13:54:31] t/full-stack.t .. 16/? # javascript console output: [
# {
#   level      => "SEVERE",
#   message    => "http://localhost:51401/asset/e27dld8f48/ws_console.js 8 WebSocket connection to 'ws://
127.0.0.1:20013/2isr2XBHNQjfdPH/ws' failed: Error in connection establishment: net::ERR_CONNECTION_REFUSED",
#   source     => "javascript",
#   timestamp  => 1588859690293,
# },
# ]
# Failed test 'no errors or warnings on javascript console, waiting for connection opened'
# at /home/squamata/project/t/lib/OpenQA/SeleniumTest.pm line 260.
# Looks like you failed 1 test of 7.
[13:54:31] t/full-stack.t .. 17/?

# Failed test 'wait until developer console becomes available'
# at t/full-stack.t line 144.
[13:54:31] t/full-stack.t .. 18/? [13:54:31] t/full-stack.t .. 19/? [13:54:31] t/full-stack.t .. 24/? [13:54:3
1] t/full-stack.t .. 25/? [13:54:31] t/full-stack.t .. 26/? [13:54:31] t/full-stack.t .. 27/? [13:54:31] t/fu
ll-stack.t .. 30/? [13:54:31] t/full-stack.t .. 31/? [13:54:31] t/full-stack.t .. 32/? [13:54:31] t/full-stack.
t .. 34/? [13:54:31] t/full-stack.t .. 36/? [13:54:31] t/full-stack.t .. 42/? [13:54:31] t/full-stack.t .. 47/
? [13:54:31] t/full-stack.t .. 50/? # Looks like you failed 1 test of 51.
[13:54:31] t/full-stack.t .. Dubious, test returned 1 (wstat 256, 0x100)
Failed 1/51 subtests
```

also we still have "RETRY=3" for full-stack.t and of course for others.

#23 - 2020-05-12 10:09 - okurz

Recent instabilities have been introduced and fixed by mkkittler, e.g. the modules details ajax route. So we can reconduct tests again and collect current state.

Trying to come back to t/05-scheduler-full.t which seems to be a problem more often lately after I added a change to prevent perl trying to access an undefined hash. The problem is that this caused a hard crash which gave an opportunity to terminate the test early. Now the test is stuck waiting for the workers which continue to run until the CI job times out so the retry on the level of the Makefile has no effect.

I prepared a change to move the complete process handling to IPC::Run to help with that but there seems to be a different underlying, scheduling related problem. In my branch "fix/scheduler3" I try to fix them. Using https://github.com/okurz/scripts/blob/master/count_fail_ratio I conducted another measurement to find out the fail ratio:

```
fails=0; for i in {1..20}; do echo "## Run $i" && env SCHEDULER_FULLSTACK=1 FULLSTACK=1 KEEP_DB=1 STABILITY_TEST=1 OPENQA_SCHEDULER_TEST_UNSTABLE_COUNT=10 OPENQA_SCHEDULER_TEST_UNRESPONSIVE_COUNT=10 DIE_ON_FAIL=1 timeout --foreground -v 60 prove -l t/05-scheduler-full.t || fails=$((fails+1)); echo "run:$i. fails: $fails. fail ratio $((fails * 100/i))"; done
```

yields run:20. fails: 11. fail ratio 55% with all fails in "one job allocated" at t/05-scheduler-full.t line 139." so I can shortcut the test to just this parts or also rework some other parts first but count the fail ratio again to ensure I have not introduced other problems as well.

<https://github.com/os-autoinst/openQA/pull/3109> for a first fix. After that – as one can see that I reached the branch name "fix/scheduler5" already – I have multiple additional changes planned, including better process handling with IPC::Run.

EDIT: 2020-05-28: Finally managed to extract one further big part out of my always-wip branch into <https://github.com/os-autoinst/openQA/pull/3123> for more use of IPC::Run and then I can continue with the next parts.

#24 - 2020-05-28 12:13 - cdywan

<https://github.com/os-autoinst/openQA/pull/3123>

#25 - 2020-05-30 09:11 - okurz

Well, I had mentioned that already in before :)

So that one was merged. Next in line:

- <https://github.com/os-autoinst/openQA/pull/3132> for various cleanup
- <https://github.com/os-autoinst/openQA/pull/3134> to fix "Connection refused" problems due to occupied ports

As expected there are still a lot of problems in 05-scheduler-full but the first problems have been fixed. Another problem I have seen:

```
[07:31:04] t/ui/26-jobs_restart.t ..... 2/?
# Failed test 'job is marked as restarted'
# at t/ui/26-jobs_restart.t line 126.
# got: 'minimalx@32bit'
# expected: 'minimalx@32bit (restarted)'
```

Bailout called. Further testing stopped: findChildElement: no such element: Unable to locate element: {"method":"xpath","selector":"./a[@title='new test']"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:103

```
FAILED--Further testing stopped: findChildElement: no such element: Unable to locate element: {"method":"xpath","selector":"./a[@title='new test']"} at /home/squamata/project/t/ui/./lib/OpenQA/SeleniumTest.pm:103
```

as well as locally in a stability run of full-stack.t:

```
Rerun 6 of 20 ...
t/full-stack.t ..... 95/? # javascript console output: [
# {
#   level    => "WARNING",
#   message  => "http://localhost:9526/asset/574a6e9f06/test_result.js 234 WebSocket connection to 'ws://localhost:9528/liveviewhandler/tests/6/developer/ws-proxy/status' failed: WebSocket is closed before the connection is established.",
#   source   => "javascript",
#   timestamp => 1590825943951,
# },
# {
#   level    => "WARNING",
#   message  => "http://localhost:9526/asset/574a6e9f06/test_result.js 234 WebSocket connection to 'ws://localhost:9528/liveviewhandler/tests/6/developer/ws-proxy/status' failed: WebSocket is closed before the connection is established.",
#   source   => "javascript",
#   timestamp => 1590825943954,
# },
# ]
t/full-stack.t ..... ok
t/05-scheduler-full.t ..... ok
```

```
t/33-developer_mode.t ..... ok
t/43-scheduling-and-worker-scalability.t .. ok
All tests successful.
Files=4, Tests=151, 290 wallclock secs ( 1.06 usr  0.07 sys + 123.64 cusr 19.83 csys = 144.60 CPU)
Result: PASS
...
152, 263 wallclock secs ( 1.02 usr  0.08 sys + 122.16 cusr 19.43 csys = 142.69 CPU)
Result: PASS
Rerun 10 of 20 ...
t/full-stack.t ..... 96/? # javascript console output: [
# {
#   level      => "WARNING",
#   message    => "http://localhost:9526/asset/574a6e9f06/test_result.js 272 WebSocket connection to 'ws://localhost:9528/liveviewhandler/tests/6/developer/ws-proxy/status' failed: WebSocket is closed before the connection is established.",
#   source     => "javascript",
#   timestamp  => 1590827059683,
# },
# ]
t/full-stack.t ..... ok
```

#26 - 2020-06-10 08:54 - okurz

<https://github.com/os-autoinst/openQA/pull/3168> for currently unstable t/ui/26-jobs_restart.t

#27 - 2020-07-10 10:39 - okurz

- Related to action #68836: t/api/14-plugin_obs_rsyc_async.t fails sporadically in OBS added

#28 - 2020-07-20 10:39 - okurz

<https://app.circleci.com/pipelines/github/os-autoinst/openQA/3638/workflows/a58744db-c263-427b-9ea7-6553f90f3fe4/jobs/34661/steps> is an example of t/43-scheduling-and-worker-scalability.t still being unstable. The flakyness is covered by the workaround with retries so far.

#29 - 2020-09-07 14:06 - cdywan

Could you give us an update on the status of this? Maybe also add a real description instead of a merged PR.

#30 - 2020-09-09 06:10 - okurz

- Description updated

sure, updated description

#31 - 2020-09-21 08:27 - okurz

Reproduced instability in t/05-scheduler-full.t which happens very seldomly, see <https://github.com/os-autoinst/openQA/pull/3403#issuecomment-695212629> showing

```
t/05-scheduler-full.t .. 3/?
# Failed test 'Allocated maximum number of jobs that could have been allocated'
# at t/05-scheduler-full.t line 241.
# got: '9'
# expected: '10'
# Looks like you failed 1 test of 1.
t/05-scheduler-full.t .. 4/?
# Failed test 'Simulation of heavy unstable load'
# at t/05-scheduler-full.t line 277.
Died at t/05-scheduler-full.t line 241.
# Tests were run but no plan was declared and done_testing() was not seen.
t/05-scheduler-full.t .. Dubious, test returned 1 (wstat 256, 0x100)
Failed 1/4 subtests
```

Reproduced with env runs=100 count_fail_ratio prove -l t/05-scheduler-full.t where it shows up exactly once (!).

Running time env runs=400 count_fail_ratio prove -l t/05-scheduler-full.t reproduced the issue 2 times, one other failure. The overall run took 215m.

Have done the same again on lord.arch:

```
$ time env runs=400 count_fail_ratio prove -l t/05-scheduler-full.
...
## count_fail_ratio: Run: 400. Fails: 5. Fail ratio 1%
real 414m38.126s
```

4 times the test ran into the timeout of 150s and only once it produced:

```
## Run 279
t/05-scheduler-full.t .. 3/?
#   Failed test 'Allocated maximum number of jobs that could have been allocated'
#   at t/05-scheduler-full.t line 241.
#       got: '9'
#     expected: '10'
# Looks like you failed 1 test of 1.
t/05-scheduler-full.t .. 4/?
#   Failed test 'Simulation of heavy unstable load'
#   at t/05-scheduler-full.t line 277.
Died at t/05-scheduler-full.t line 241.
# Tests were run but no plan was declared and done_testing() was not seen.
t/05-scheduler-full.t .. Dubious, test returned 1 (wstat 256, 0x100)
Failed 1/4 subtests
```

Test Summary Report

So around 1/400 runs reproduces the issue.

Created specific ticket [#72085](#) for this.

#32 - 2020-09-24 05:13 - okurz

- Related to action [#71554](#): *unstable/flaky/sporadic t/full-stack.t test failing in script waits on CircleCI added*

#33 - 2020-09-29 14:44 - okurz

- Copied to action [#72085](#): *flaky/unstable t/05-scheduler-full.t, Failed test 'Allocated maximum number of jobs that could have been allocated' added*

#34 - 2020-09-29 14:44 - okurz

- Target version changed from *Current Sprint* to *Ready*

#35 - 2020-11-03 10:46 - okurz

- Related to action [#76900](#): *unstable/flaky/sporadic t/full-stack.t test failing in CircleCI "worker did not propagate URL for os-autoinst cmd srv within 1 minute" added*

#36 - 2020-11-03 10:47 - okurz

- Status changed from *Feedback* to *Blocked*

see related tasks

#37 - 2020-11-25 12:04 - okurz

- Estimated time set to *39719.00 h*

#38 - 2020-11-25 12:10 - okurz

- Estimated time deleted (*39719.00 h*)

#39 - 2021-02-09 12:49 - okurz

- Copied to action [#88496](#): *openQA "t" tests time out (again) and take long added*

#40 - 2021-02-11 19:43 - okurz

- Description updated

- Status changed from *Blocked* to *Resolved*

Hm, I think we should have kept the original focus for this test and not try to extend it endlessly. I should have known that there will likely always be some unstable tests :/ The original problem in `t/fullstack.t` is long fixed.