

[saga][epic] Detection of "known failures" for stable tests, easy test results review and easy tracking of known issues

2018-08-14 13:15 - okurz

Status:	Blocked	Start date:	2017-01-23
Priority:	High	Due date:	2021-06-29
Assignee:	okurz	% Done:	59%
Category:	Feature requests	Estimated time:	0.00 hour
Target version:	Ready		
Difficulty:			

Description

User Story

As a reviewer of failed openQA tests I want known failures of jobs regardless of the error source to be marked as such automatically to not waste time on investigating known failures

Acceptance criteria

- **AC1:** If a job fails for any reason that is "known" already in the context of the current openQA instance no further "test review" effort is needed by human reviewers

Suggestions

- Provide a mechanism to match on regex in serial0.txt (as provided by existing "serial exception catching"-feature) based on patterns defined in the test distribution
- Same for autoinst-log.txt
- Provide patterns defined in os-autoinst for backend specific stuff, e.g. the "key event queue full"-thingy -> look for that string in os-autoinst for existing code to handle that
- Same as above but patterns defined in instance specific configuration, e.g. workers.ini (managed by salt for SLE)
- Maybe the same based on needles? But maybe the current approach using the "workaround" property and soft-fail needles to be always preferred is already good enough :)
- It might be necessary to re-define "soft-fail" as "known issue" and nothing more so that we can use the "known failure" detection to set a job to soft-failed referencing the known issue, immediately aborting the further execution of a job to prevent it failing at a sporadic later step which would pose the need to provide openQA comments to provide a label

Further details

Definitions:

- "known" means that a certain symptom of a test failure has been described with e.g. a matching pattern in either a test distribution, os-autoinst or maybe openQA itself as for the later mentioned jenkins plugin
- "test review" means what we currently do in openSUSE or SLE by providing job labels with issue references in openQA comments which are carried over – which so far only works within individual scenarios

See <https://wiki.jenkins.io/display/JENKINS/Build+Failure+Analyzer> for an example. This jenkins plugin uses a "knowledge base" with jenkins instance global "known failures" defined with description and pattern matching, e.g. on "build log parsing", to mark failures as known when any log content matches existing patterns

Subtasks:

action # 16180: Better log viewer	New
coordination # 19720: [epic] Simplify investigation of job failures	Blocked
action # 61103: Use CodeMirror to render diffs in the Investigation tab	Rejected
action # 69085: Make "last good" a link to a job instead of plain job ID	Resolved
action # 69088: Present changes between packages on openQA worker machines in "investig...	Resolved
action # 69319: Surface relevant versions used by job worker	New
coordination # 91518: [epic] Provide 'first bad' vs. 'last good' difference in investig...	Blocked
action # 91521: link to "first bad" in investigation tab	Resolved

action # 92188: test reviewers are pointed to the "first bad vs. last good" comparison ...	Resolved
action # 91527: Cleanup logging in autoinst-log.txt	Resolved
action # 91878: Improve git log entries in failed test investigation	Workable
action # 92731: clickable git log entries in investigation tab	Workable
action # 92746: Log viewer in openQA webUI with color parsing	Resolved
action # 93940: text thumbnail preview feels inconsistent to other screenshots	Workable
action # 38621: [functional][y] test fails in welcome - "Module is not signed with expe...	Resolved
action # 46988: [qe-core][functional] Detect known bugs from system journal	New
action # 60560: Self-investigate potential reasons for failures in openQA	Resolved
coordination # 62420: [epic] Distinguish all types of incompletes	Blocked
action # 45062: Better visualization of incompletes - show module in which incomplete h...	Resolved
action # 59926: test incompletes in middle of execution with auto_review:"Unexpected en...	New
coordination # 61922: [epic] Incomplete jobs with no logs at all	Resolved
action # 62984: Fix problem with job-worker assignment resulting in API errors	Resolved
action # 63718: incomplete reason with just "quit"/"died" could provide more information	Resolved
action # 64854: qemu-img error message is incorrectly tried to be parsed as JSON auto_r...	Resolved
action # 64857: Put single-line error messages into incomplete reason for "died"	Resolved
action # 64884: Distinguish test contributor errors from unexpected backend crashes	Resolved
action # 64917: auto_review:"(?s)qemu-img.*runcmd.*failed with exit code 1" sometimes b...	Resolved
action # 66066: incomplete with reason "died: terminated prematurely" but log shows err...	Rejected
action # 67000: Job incompletes due to malformed worker cache database disk image with ...	Resolved
action # 69448: test incompletes with auto_review:"(?s)was downloaded by.*details.*unav...	Workable
coordination # 69451: [epic] test incompletes with "(?s)Download.*successful.*Failed to...	New
action # 69553: job incompletes with "Failed to rsync tests: exit code 10":retry, impro...	Resolved
action # 69691: Improve incomplete output for qemu related problems, e.g. auto_review:"...	Workable
action # 71185: job incompletes with auto_review:"setup failure: Cache service status e...	Resolved
action # 71188: job incomplete with auto_review:"backend died: QEMU exited unexpectedly...	Workable
action # 71227: job incompletes with auto_review:"backend died: 'current_console' is no...	Workable
action # 71827: test incompletes with auto_review:"(?s)Failed to download.*Asset was pr...	Resolved
action # 73273: job incompletes with auto_review:"setup failure: Cache service status e...	Workable
action # 73282: auto_review:"setup failure: Cache service status error from API: Minion...	Workable
action # 73285: test incompletes with auto_review:"(?s)Download of.*processed[^.]*Fail...	Resolved
action # 73288: auto_review:"setup failure: Cache service status error from API: Minion...	Workable
action # 73294: auto_review:"isotovideo died: needles_dir not found" should be 'tests d...	Workable
action # 73339: auto_review:"setup failure: Cache service status error from API: Minion...	Resolved
action # 73369: Job incompletes with auto_review:"(?s)backend died: runcmd .*qemu-img c...	Workable
action # 73375: Job incompletes with reason auto_review:"(?m)api failure\$" (and no furt...	Workable
action # 73396: job incompletes with auto_review:"setup failure: Failed to rsync tests:...	Resolved
action # 73525: Job incompletes with auto_review:"backend died: unexpected end of data ...	New
action # 75388: Explicit error feedback to test reviewers on wrong test API usage	Workable
action # 78055: job incomplete exiting prematurely before reaching needle check timeout...	New
action # 78169: after osd-deploy 2020-11-18 incompletes with auto_review:"Cache service...	Resolved
action # 80106: corrupted worker cache sqlite: Enlarge systemd service kill timeout tem...	Resolved
action # 80118: test incompletes with auto_review:"(?s)Failed to download.*Asset was pr...	Resolved
action # 80226: job incomplete with autoinst-log.txt ending just in the middle	Workable
action # 80334: job incompletes with auto_review:"(?s)terminated prematurely with corru...	Resolved
action # 80356: incompletes with auto_review:"Cache service.*error: Connection refused"...	Workable
action # 80408: revert longer timeout override for openQA services as we could not see ...	Resolved
action # 80778: job incompletes with "Virtio terminal and svirt serial terminal do not ...	Workable
action # 89614: openqa workers on `ip-172-25-5-39` fails with no clue on failure	Resolved
action # 90974: Make it obvious if qemu gets terminated unexpectedly due to out-of-memory	Feedback
action # 63065: [gsoc] dynamic detection of error conditions from test results	New
action # 64935: Provide test results of "related results" on job details	Workable

action # 35017: Show more test failure details with the "test failed" text thumbnail po...	Workable
coordination # 41057: [epic] Make reviewing results easier	Blocked
action # 91601: Add "return to top" button on openQA pages, e.g. job details, index, gr...	Resolved
action # 91602: Inform users about keyboard shortcuts in openQA	Workable
action # 62441: openqa-worker systemd service can timeout when stopping	Workable
action # 75232: error message when worker has no network (yet): Unable to serialize fat...	Resolved
coordination # 77899: [epic] Extend "auto-review" for failed jobs as well	Resolved
action # 80414: [proof-of-concept] Extend "auto-review" for failed jobs as well, start ...	Resolved
action # 80418: [learning] Fix parse errors in "openqa-investigate" "parse error: Inval...	Resolved
action # 80806: Extend "auto-review" for failed jobs as well - Generalize openqa-monito...	Resolved
action # 80808: Extend "auto-review" for failed jobs as well - enable same as on o3 bu...	Resolved
action # 77944: Run "auto-review" more often but alarm less	Resolved
action # 80264: multimachine tests unable to get vars from its pair job	Resolved
action # 80412: tests fail with auto_review:"(?s)version is 4\.\6\.\1606298538\.\191b5988....	Resolved
action # 80772: [jeos] auto_review:"(?s)GENERAL_HW_FLASH_CMD.*No space left on device":...	Resolved
action # 80774: [jeos] auto_review:"(?s)GENERAL_HW_FLASH_CMD.*No route to host":retry i...	Resolved
coordination # 80828: [epic] Trigger 'auto-review' and 'openqa-investigate' from within...	Resolved
action # 80736: Trigger 'auto-review' from within openQA when jobs incomplete (or fail)...	Resolved
action # 80826: Trigger 'auto-review' from within openQA when jobs incomplete on osd as...	Resolved
action # 80830: Trigger 'openqa-investigate' from within openQA when jobs fail on o3	Resolved
action # 81206: Trigger 'openqa-investigate' from within openQA when jobs fail on osd	Resolved
action # 81859: openqa-investigate triggers incomplete sets for multi-machine scenarios	Resolved
coordination # 88229: [epic] Prevent unintended test coverage decrease	Blocked
action # 88127: [tools][qem] Test coverage DB for maintenance updates	Resolved
action # 88485: [teregen] Fetch and store coverage info for each incident	Resolved
action # 90401: [teregen] Integrate coverage information in a presentable way into test...	Resolved
action # 90404: [teregen] Update TeReGen for deployment on qam2	Resolved
action # 88536: Find out differences in openQA test coverage with metabase	Resolved
action # 91509: [tools] Easy way to check and compare coverage in multiple openQA insta...	New
action # 91656: [qe-core] os-autoinst-distri-opensuse YAML schedule file comparison	New

Related issues:

Related to openQA Project - action #13242: WDYT: For every job that does not ...	Rejected	2016-11-25
Related to openQA Project - coordination #13812: [epic][dashboard] openQA Das...	New	2017-01-10
Related to openQA Tests - action #42446: [qe-core][functional] many opensuse ...	New	2018-10-13
Related to openQA Project - action #40382: Make "ignored" issues more promine...	Workable	2018-08-29
Related to openQA Tests - action #43784: [functional][y][sporadic] test fails...	Resolved	2018-11-14
Related to openQA Project - action #57452: Automatic summary of failures	Rejected	2019-09-27
Blocked by openQA Project - action #45011: Allow detection of known failures ...	Workable	2018-12-11

History

#1 - 2018-08-14 13:15 - okurz

- Related to action #13242: WDYT: For every job that does not have a label or bugref, retrigger some times to see if it's sporadic. Like rescheduling on incomplete but on failed added

#2 - 2018-08-14 13:15 - okurz

- Related to action #38621: [functional][y] test fails in welcome - "Module is not signed with expected PKCS#7 message" (bsc#1093659) - Use serial exception catching feature from openQA to make sure the jobs reference the bug, e.g. as label added

#3 - 2018-08-15 06:50 - okurz

- Related to coordination #13812: [epic][dashboard] openQA Dashboard ideas added

#4 - 2018-08-15 07:57 - okurz

- Related to deleted (action #38621: [functional][y] test fails in welcome - "Module is not signed with expected PKCS#7 message" (bsc#1093659) - Use serial exception catching feature from openQA to make sure the jobs reference the bug, e.g. as label)

#5 - 2018-08-22 11:41 - nicksinger

Another idea which could be checked/better reported to the user:

If a crucial component in the "os-autoinst-chain" fails (e.g. xterm for ipmi jobs), openQA could easily report this earlier. As it is right now, the job stalls (hangs as "running") but only shows a black screen. Example: <https://openqa.suse.de/tests/1970948> (look for "PermissionError" in the osautoinst-log.txt)

#6 - 2018-10-07 08:25 - coolo

- Target version set to future

IMO this is best handled by an automated review from outside. The problem is not so much the detecting the issue, but how to handle it. For some projects/objects you would do a retrigger, for others you would prefer defining a label.

#7 - 2018-10-07 08:31 - okurz

"outside", yes, I agree. Should be outside what is currently defined as "openQA" but it could be that we still call it "the openQA ecosystem" so I guess this issue tracker is still best suited. Some parts we have already covered with the proof-of-concept of detecting known failures in the serial port output.

#8 - 2018-10-07 09:07 - coolo

I don't disagree with the issue tracker - I just don't want a High priority epic in my 'to be sorted' list

#9 - 2018-10-13 22:44 - okurz

- Related to action #42446: [qe-core][functional] many opensuse tests fail in desktop_runner or gimp or other modules in what I think is boo#1105691 – can we detect this bug from the journal and track as soft-fail? added

#10 - 2018-10-13 22:48 - okurz

- Subject changed from [epic] Detect "known failures" and mark jobs as such to [functional][y][u][epic] Detect "known failures" and mark jobs as such

Trying to bring it forward with help of QSF again...

#11 - 2018-10-14 18:49 - okurz

- Related to action #27004: [opensuse][sle][functional][yast][y][hard] yast2 gui modules fail to start in the defined time frame added

#12 - 2018-10-14 18:49 - okurz

- Related to deleted (action #27004: [opensuse][sle][functional][yast][y][hard] yast2 gui modules fail to start in the defined time frame)

#13 - 2018-10-14 18:49 - okurz

- Blocks action #27004: [opensuse][sle][functional][yast][y][hard] yast2 gui modules fail to start in the defined time frame added

#14 - 2018-10-18 09:25 - okurz

- Related to action #40382: Make "ignored" issues more prominent (was: create new state "ignored") added

#15 - 2018-11-13 14:19 - okurz

<https://github.com/os-autoinst/os-autoinst/pull/1052> to "Add option to override status of test modules with soft-fail"

#16 - 2018-11-14 07:50 - okurz

- Status changed from New to Feedback

- Assignee set to okurz

#17 - 2018-11-15 16:03 - okurz

The feature is not working as intended as in <https://github.com/os-autoinst/os-autoinst/blob/master/basetest.pm#L286> we overwrite the result again. I am trying to simply remove that method :)

-> <https://github.com/os-autoinst/os-autoinst/pull/1062>

Also presented my idea to riafarov and we identified one problematic scenario: What if we force the status of a parent job to "softfail"? For now openQA would still trigger the downstream jobs which then most likely should fail because a module in the parent job failed, in the worst case even making the downstream jobs incomplete because the HDD image was never published properly. We should avoid this though.

#18 - 2018-12-04 10:32 - okurz

- Related to action #43784: [functional][y][sporadic] test fails in yast2_snapper now reproducibly not exiting the "show differences" screen added

#19 - 2018-12-11 16:52 - szarate

- Related to action #45011: Allow detection of known failures at the autoinst-log.txt added

#20 - 2018-12-11 16:53 - szarate

I see that one of the suggestions on this ticket was exactly what poo#45011 is about :)

#21 - 2019-01-11 10:06 - agraul

- Related to deleted (action #45011: Allow detection of known failures at the autoinst-log.txt)

#22 - 2019-01-11 10:06 - agraul

- Blocked by action #45011: Allow detection of known failures at the autoinst-log.txt added

#23 - 2019-01-11 10:06 - agraul

- Status changed from Feedback to Blocked

[#45011](#)

#24 - 2019-02-05 06:34 - okurz

- Due date changed from 2018-08-28 to 2019-03-12

due to changes in a related task

#25 - 2019-03-12 16:24 - okurz

- Due date changed from 2019-03-12 to 2019-06-30

due to changes in a related task

#26 - 2019-05-22 10:19 - okurz

- Assignee changed from okurz to riafarov

Move to new QSF-y PO after I moved to the "tools"-team. I mainly checked the subject line so in individual instances you might not agree to take it over completely into QSF-y. Feel free to reassign to me or someone else in this case. Thanks.

#27 - 2019-05-23 08:10 - riafarov

- Blocks deleted (action #27004: [opensuse][sle][functional][yast][y][hard] yast2 gui modules fail to start in the defined time frame)

#28 - 2019-07-02 10:59 - riafarov

- Due date changed from 2019-06-30 to 2019-08-06

due to changes in a related task

#29 - 2019-08-06 08:50 - riafarov

- Due date changed from 2019-08-06 to 2019-12-31

due to changes in a related task

#30 - 2019-09-27 11:32 - okurz

- Related to action #57452: Automatic summary of failures added

#31 - 2019-12-03 15:50 - okurz

Using https://github.com/os-autoinst/scripts/blob/master/monitor-openqa_job and <https://github.com/os-autoinst/scripts/blob/master/openqa-label-known-issues> I setup a gitlab CI pipeline in <https://gitlab.suse.de/openqa/auto-review/> that automatically labels (and restarts) incompletes for which we know the reasons. The approach could also be extended to cover not only incompletes.

#32 - 2019-12-17 16:14 - okurz

- Related to coordination #19720: [epic] Simplify investigation of job failures added

#33 - 2020-01-09 08:47 - riafarov

- Assignee changed from riafarov to okurz

As it's mainly tools team working on this epic, [okurz](#) I will set you as an assignee to track the progress. Feel free to change it, I rely on your expertise to set more suitable person if it's not you. Thanks!

#34 - 2020-01-09 09:19 - okurz

- Subject changed from [functional][y][u][epic] Detect "known failures" and mark jobs as such to [epic] Detect "known failures" and mark jobs as such

that's ok, it's me :)

There is currently only one subtask open [#46988](#) on QSF-u though.

#35 - 2020-01-13 11:33 - okurz

- Due date changed from 2019-12-31 to 2020-12-31

due to changes in a related task

#36 - 2020-03-05 14:49 - okurz

- Subject changed from [epic] Detect "known failures" and mark jobs as such to [saga] Detect "known failures" and mark jobs as such

#37 - 2020-03-05 14:54 - okurz

- Subject changed from [saga] Detect "known failures" and mark jobs as such to [saga][epic] Detect "known failures" and mark jobs as such

#38 - 2020-05-04 11:30 - SLindoMansilla

- Due date changed from 2020-12-31 to 2020-03-27

due to changes in a related task: [#46988](#)

#39 - 2020-07-03 19:21 - okurz

- Due date changed from 2020-06-09 to 2020-03-27

due to changes in a related task: [#62420](#)

#40 - 2020-07-03 20:07 - okurz

- Due date changed from 2018-08-28 to 2020-03-27

due to changes in a related task: [#38621](#)

#41 - 2020-10-12 13:31 - szarate

- Tracker changed from action to coordination

- Status changed from Blocked to New

#42 - 2020-10-12 13:46 - szarate

See for the reason of tracker change: <http://mailman.suse.de/mailman/private/qa-sle/2020-October/002722.html>

#43 - 2020-10-13 11:53 - okurz

- Status changed from New to Blocked

- Target version changed from future to Ready

Discussed the topic of "auto-review" with SUSE QA Tools team and the general opinion was that this epic is interesting to follow up with so putting it to the backlog now.

#44 - 2020-11-22 09:25 - okurz

- Subject changed from [saga][epic] Detect "known failures" and mark jobs as such to [saga][epic] Detect "known failures" and mark jobs as such to make tests more stable, reviewing test results and tracking known issues easier

#45 - 2020-12-03 10:42 - cdywan

Once again wondering: where's the due date coming from? It's not visible. Do we need to go through every single ticket again to check?

#46 - 2020-12-03 20:42 - okurz

Maybe the API helps to find that easily but in this case it's [#80264](#)

#47 - 2021-04-28 21:19 - okurz

- Subject changed from [saga][epic] Detect "known failures" and mark jobs as such to make tests more stable, reviewing test results and tracking known issues easier to [saga][epic] Detection of "known failures" for stable tests, easy test results review and easy tracking of known issues