Multiple failures due to network issues

I will use this ticket to collect the different errors I observe in our tests (at least for QE-C squad) that fail due to network issues. Normally a restart helps to get the job green again (in case we need it green), but this is not the ideal solution.

The idea of this ticket is to collect more potential issues caught by reviewers and propose solutions for some of them, in the code (retry same command several times might help) or in the infra side.

There is an example for each error I found, but from my experience reviewing jobs every day, these failures happen multiple times a day and randomly (difficult to predict).

1) SUSEConnect timeouts
   Test died: command 'SUSEConnect -p sle-module-containers/$(VERSION_ID)/$(CPU)' timed out at /usr/lib/os-autoinst/testapi.pm line 1039.
   Or https://openqa.suse.de/tests/8193554#step/suseconnect_scc/8
   Test died: command 'SUSEConnect -r $regcode' timed out at /usr/lib/os-autoinst/testapi.pm line 950.

2) updates.suse.com not reachable
   Test died: command 'SUSEConnect -p sle-module-containers/$(VERSION_ID)/$(CPU)' timed out at /usr/lib/os-autoinst/testapi.pm line 1039.
   Or https://openqa.suse.de/tests/8193554#step/suseconnect_scc/8
   Test died: command 'SUSEConnect -r $regcode' timed out at /usr/lib/os-autoinst/testapi.pm line 950.

3) SCC timeouts
   Test died: command 'SUSEConnect -p sle-module-containers/$(VERSION_ID)/$(CPU)' timed out at /usr/lib/os-autoinst/testapi.pm line 1039.
   Or https://openqa.suse.de/tests/8193554#step/suseconnect_scc/8
   Test died: command 'SUSEConnect -r $regcode' timed out at /usr/lib/os-autoinst/testapi.pm line 950.

4) zypper ref timeout or error
   Test died: command 'SUSEConnect -p sle-module-containers/$(VERSION_ID)/$(CPU)' timed out at /usr/lib/os-autoinst/testapi.pm line 1039.
   Or https://openqa.suse.de/tests/8193554#step/suseconnect_scc/8
   Test died: command 'SUSEConnect -r $regcode' timed out at /usr/lib/os-autoinst/testapi.pm line 950.
Acceptance criteria

- AC1: All existing subtasks are resolved, no additional work needed on top

Subtasks:

<table>
<thead>
<tr>
<th>Action ID</th>
<th>Subtask Description</th>
<th>Status</th>
<th>Resolution Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>99345</td>
<td>[tools][qem] Incomplete test runs on s390x with auto_review:“backend di...</td>
<td>Resolved</td>
<td></td>
</tr>
<tr>
<td>108266</td>
<td>grenache: script_run() commands randomly time out since server room move</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>108845</td>
<td>Network performance problems, DNS, DHCP, within SUSE QA network auto_r...</td>
<td>Resolved</td>
<td></td>
</tr>
<tr>
<td>108872</td>
<td>Outdated information on openqaw5-xen <a href="https://racktables.suse.de/index">https://racktables.suse.de/index</a>....</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>108896</td>
<td>[ppc64le] auto_review:“(?s)Size of.&quot;differ, expected.&quot;but downloaded....</td>
<td>Resolved</td>
<td></td>
</tr>
<tr>
<td>109241</td>
<td>[tools] Performance issues in some s390 workers</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>109253</td>
<td>Add monitoring for SUSE QA network infrastructure size:M</td>
<td>Resolved</td>
<td></td>
</tr>
<tr>
<td>108668</td>
<td>Failed systemd services al...</td>
<td>Rejected</td>
<td></td>
</tr>
<tr>
<td>113528</td>
<td>[qe-core] test fails in bootloader_...</td>
<td>Resolved</td>
<td></td>
</tr>
</tbody>
</table>

Related issues:

- Related to openQA Tests - action #107635: [qem][y] test fails in installation       | New                                      | 2022-02-25   |
- Related to openQA Infrastructure - action #108668: Failed systemd services al...   | Rejected                                 | 2022-03-21   |
- Related to openQA Tests - action #113528: [qe-core] test fails in bootloader_...   | Resolved                                 | 2022-07-18   |
- Related to openQA Infrastructure - action #113716: [qe-core] proxy-scc is down     | Resolved                                 | 2022-07-19   |

History

#1 - 2022-02-18 09:17 - jlausuch

More examples of failures/timeouts activating a module via SUSEConnect:

- https://openqa.suse.de/tests/8189772#step/docker_compose/5
- https://openqa.suse.de/tests/8189767#step/docker_compose/5

#2 - 2022-02-18 11:03 - jlausuch

I have created this proposal for adding a product/module with suseconnect.
https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14283

#4 - 2022-02-19 09:09 - jlausuch

- Description updated

#6 - 2022-02-21 09:48 - jlausuch

- File scc_timeout.png added

2 more issues due to SCC not reachable:
https://openqa.suse.de/tests/8201102#step/suseconnect_scc/6
https://openqa.suse.de/tests/8201100#step/suseconnect_scc/6

scc_timeout.png

#7 - 2022-02-21 10:03 - jlausuch

- File Screenshot 2022-02-21 at 11.03.33.png added

And another one with updates.suse.com not reachable:
https://openqa.suse.de/tests/8201099#step/glibc_locale/47

2022-09-04
New PR to create a new retry wrapper for validate_script_output method.
https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14295

This will help with issues with validate_script_output that are prone to timeout if network is slow.

#9 - 2022-02-24 08:10 - dzedro
Few examples of SCC/network failures, it's nothing rare.
https://openqa.suse.de/tests/8219630#step/zypper_patch/20
https://openqa.suse.de/tests/8198732#step/scc_registration/16

#10 - 2022-02-25 09:40 - jlausuch
dzedro wrote:

Few examples of SCC/network failures, it's nothing rare.
https://openqa.suse.de/tests/8219630#step/zypper_patch/20

For this, we can make a script_retry, similar to what I did here: https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14327
https://openqa.suse.de/tests/8198732#step/scc_registration/16

This is more tricky, as it's installation UI, we can't use any in-code retry here... (it would need needle handling for error cases)

#11 - 2022-02-25 12:08 - okurz
- Related to action #107635: [qem][y] test fails in installation added

#12 - 2022-02-25 13:33 - okurz
As suggested over other channels might have gone lost: I suggest to get in contact with the SCC team and EngInfra-Team, e.g. create bugs in their according issue trackers and link here or ping them over chat or email and invite them to brainstorm and investigate on this issue. We don't know where the problem is but I am sure together with experts from other teams we have enough brain power to solve this riddle. It could be that there is a problem within the product that we or you test and then of course we should handle this as appropriate product problems or regressions that should be fixed because then customers will also be affected. It could be a problem in the user facing infrastructure especially for updates.suse.com on a CDN where customers would or could also be affected. Then as well we should not just accept this issue but look into it and try to solve it together.

#13 - 2022-02-28 18:11 - jlausuch
okurz wrote:

As suggested over other channels might have gone lost: I suggest to get in contact with the SCC team and EngInfra-Team, e.g. create bugs in their according issue trackers and link here or ping them over chat or email and invite them to brainstorm and investigate on this issue. We don't know where the problem is but I am sure together with experts from other teams we have enough brain power to solve this riddle. It could be that there is a problem within the product that we or you test and then of course we should handle this as appropriate product problems or regressions that should be fixed because then customers will also be affected. It could be a problem in the user facing infrastructure especially for updates.suse.com on a CDN where customers would or could also be affected. Then as well we should not just accept this issue but look into it and try to solve it together.

Yes, that would be ideal, but first we would need to collect proofs and really have specific tests for SCC/updates.suse.de/etc that tests the connectivity (I think there is a ticket about it) and collect some statistics to see how that behaves over the time. Maybe we can recognize a pattern when this hiccups happen, or maybe we don't. I think this is a quite complex scenario to define. I personally don't have the time to drive this initiative forward.

#14 - 2022-02-28 18:13 - jlausuch
Again: https://openqa.suse.de/tests/8242917#step/image_docker/705

> docker run --entrypoint /usr/lib/zypp/plugins/services/container-suseconnect-zypp -i zypper_docker-derived

2022/02/28 13:53:28 Installed product: SLES-12.4-x86_64
2022/02/28 13:53:28 Registration server set to https://scc.suse.com

Hopefully, these type of failures will be workarounded by https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14295

#15 - 2022-02-28 18:26 - jlausuch
And more:
https://openqa.suse.de/tests/8243068#step/registry/156
> SUSEConnect --status-text
> EOT_L_NVX
L_NVX-0-
# echo L_NVX; bash -o pipefail /tmp/scriptL_NVX.sh ; echo SCRIPT_FINISHED_L_NVX-?-
L_NVX
Error: Cannot parse response from server

#16 - 2022-02-28 21:40 - jlausuch
An interesting one, even failed after 3 retries... and this has nothing to do with our scc or any suse.de domain:  
https://openqa.suse.de/tests/8245957#step/docker_3rd_party_images/825

# timeout 420 docker pull registry.access.redhat.com/ubi7/ubi-init; echo 2n4wD-?-
Using default tag: latest
Error response from daemon: error parsing HTTP 404 response body: invalid character 'N' looking for beginning of value: "Not found"
2n4wD-1-
# timeout 420 docker pull registry.access.redhat.com/ubi7/ubi-init; echo 2n4wD-?-
Using default tag: latest
Error response from daemon: error parsing HTTP 404 response body: invalid character 'N' looking for beginning of value: "Not found"
2n4wD-1-
# timeout 420 docker pull registry.access.redhat.com/ubi7/ubi-init; echo 2n4wD-?-
Using default tag: latest
Error response from daemon: error parsing HTTP 404 response body: invalid character 'N' looking for beginning of value: "Not found"
2n4wD-1-

I guess this time we can blame RH :) 

#17 - 2022-03-01 14:02 - pdostal
jlausuch wrote:

dzedro wrote:

Few examples of SCC/network failures, it's nothing rare.  
https://openqa.suse.de/tests/8219630#step/zypper_patch/20

For this, we can make a script_retry, similar to what I did here:  
https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14327

https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14383

#18 - 2022-03-01 14:07 - pdostal
- File canvas.png added

jlausuch wrote:

dzedro wrote:

https://openqa.suse.de/tests/8198732#step/scc_registration/16

This is more tricky, as it's installation UI, we can't use any in-code retry here... (it would need needle handling for error cases)

YaST2 Installation - Connection to registration server failed. (I'm just documenting this so it's not cleaned by openQA).

#20 - 2022-03-02 13:51 - jlausuch
jlausuch wrote:

I have created this proposal for adding a product/module with suseconnect.  
https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14283

Follow-up:  
https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14402

#21 - 2022-03-03 07:51 - jlausuch
New one today in parent job where many jobs depend on...
Installation:  
https://openqa.suse.de/tests/8261955#step/scc_registration/31
#22 - 2022-03-03 08:44 - jlausuch
To keep record:
https://openqa.suse.de/tests/8261596#step/image_docker/1491

# timeout 600 docker exec refreshed zypper -nv ref; echo Uyq31-$?
Entering non-interactive mode.
Verbosity: 2
Initializing Target
Refreshing service 'container-suseconnect-zypp'.
Problem retrieving the repository index file for service 'container-suseconnect-zypp':
[container-suseconnect-zypp|file:/usr/lib/zypp/plugins/services/container-suseconnect-zypp]
Warning: Skipping service 'container-suseconnect-zypp' because of the above error.
Specified repositories:
Warning: There are no enabled repositories defined.
Use 'zypper addrepo' or 'zypper modifyrepo' commands to add or enable repositories.

This failed even after some 3 retries, and the consequent modules also failed with similar connectivity issues:

```bash
> SUSEConnect --status-text
> EOT_L_NVX
L_NVX-0-
# echo L_NVX; bash -oe pipefail /tmp/scriptL_NVX.sh ; echo SCRIPT_FINISHEDL_NVX-$?
SUSEConnect error: SocketError: getaddrinfo: Temporary failure in name resolution
```

which could be partially avoided with https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14409

#23 - 2022-03-03 09:16 - dzedro
I guess it's time to fix the real network or SCC issue.

#24 - 2022-03-03 09:18 - jlausuch

dzedro wrote:

I guess it's time to fix the real network or SCC issue.

+1

We just need a driver for that :)

#26 - 2022-03-03 10:11 - jlausuch

https://openqa.suse.de/tests/8262854#step/image_docker/921

```bash
# docker run --rm -ti registry.suse.com/suse/sle15:15.2 zypper ls | grep 'Usage:'; echo GcdHR-$?
GcdHR-1-
# cat > /tmp/scriptTBhqL.sh << 'EOT_TBhqL'; echo TBhqL-$?
> docker run -i --entrypoint '' registry.suse.com/suse/sle15:15.2 zypper lr -s
> EOT_TBhqL
TBhqL-0-
# echo TBhqL; bash -oe pipefail /tmp/scriptTBhqL.sh ; echo SCRIPT_FINISHEDTBhqL-$?
SUSEConnect error: SocketError: getaddrinfo: Temporary failure in name resolution
```

#27 - 2022-03-04 07:47 - jlausuch

Today again parent installation jobs failing in SCC registration step:
https://openqa.suse.de/tests/8269572#step/scc_registration/29
https://openqa.suse.de/tests/8269587#step/scc_registration/36

#28 - 2022-03-04 07:51 - jlausuch

SLE Micro timeout in transactional-update migration:
https://openqa.suse.de/tests/8268653#step/zypper_migration/9

2022-09-04
Calling zypper migration --no-snapshots --no-selfupdate
2022-03-04 01:25:26 tukit 3.4.0 started
2022-03-04 01:25:26 Options: call 5 zypper migration --no-snapshots --no-selfupdate
2022-03-04 01:25:26 Executing 'zypper migration --no-snapshots --no-selfupdate':

Executing 'zypper refresh'

Repository 'SUSE-MicroOS-5.0-Pool' is up to date.
Repository 'SUSE-MicroOS-5.0-Updates' is up to date.
Repository 'TEST_0' is up to date.
Repository 'TEST_1' is up to date.
Repository 'TEST_10' is up to date.
Repository 'TEST_11' is up to date.
Repository 'TEST_12' is up to date.
Repository 'TEST_2' is up to date.
Repository 'TEST_3' is up to date.
Repository 'TEST_4' is up to date.
Repository 'TEST_5' is up to date.
Repository 'TEST_6' is up to date.
Repository 'TEST_7' is up to date.
Repository 'TEST_8' is up to date.
Repository 'TEST_9' is up to date.
All repositories have been refreshed.
Can't determine the list of installed products: JSON::ParserError: 765: unexpected token at '<html>
<head><title>504 Gateway Time-out</title></head>
<body>
<center><h1>504 Gateway Time-out</h1></center>
</body>
</html>'

'/usr/lib/zypper/commands/zypper-migration' exited with status 1
2022-03-04 01:25:42 Application returned with exit status 1.

#29 - 2022-03-04 07:54 - jlausuch
Timeouts on some container command running zypper lr:
https://openqa.suse.de/tests/8269339#step/image_docker/919
https://openqa.suse.de/tests/8269313#step/image_docker/162
https://openqa.suse.de/tests/8269116#step/image_docker/929

> docker run --entrypoint '' registry.suse.com/suse/sles12sp5 zypper lr -s
> EOT_sed1h
sed1h-0-
# echo sed1h; bash -oe pipefail /tmp/scriptsed1h.sh; echo SCRIPT_FINISHEDsed1h-5?-sed1h
Refreshing service 'container-suseconnect-zypp'.
Problem retrieving the repository index file for service 'container-suseconnect-zypp':
[container-suseconnect-zypp|file:/usr/lib/zypp/plugins/services/container-suseconnect-zypp]
Warning: Skipping service 'container-suseconnect-zypp' because of the above error.
Warning: No repositories defined.
Use the 'zypper addrepo' command to add one or more repositories.
SCRIPT_FINISHEDsed1h-6-

This can be partially stabilized with retries: https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14420

#30 - 2022-03-04 07:59 - jlausuch
btrfs_autocompletion:
Timeout exceeded when accessing
'https://scc.suse.com/access/services/1924/repo/repoindex.xml?credentials=SUSE_Linux_Enterprise_Server_15_SP2_x86_64'.
https://openqa.suse.de/tests/8268555#step/btrfs_autocompletion/11
https://openqa.suse.de/tests/8268561#step/btrfs_autocompletion/11

#31 - 2022-03-07 08:04 - jlausuch
Today's failures:
• installation job, scc timeout: https://openqa.suse.de/tests/8280236#step/scc_registration/15

• btrfs_autocompletion:
  File '/media.1/media' not found on medium 'http://dist.suse.de/ibs/SUSE:/Maintenance:/22627/SUSE_Updates_SLE-SERVER_12-SP5_x86_64/'
  https://openqa.suse.de/tests/8279343#step/btrfs_autocompletion/11
  https://openqa.suse.de/tests/8279356#step/btrfs_autocompletion/11
  https://openqa.suse.de/tests/8279355#step/btrfs_autocompletion/10

• docker|podman build timeouts due to container-suseconnect-zypp:
  https://openqa.suse.de/tests/8280575#step/image_docker/771
  https://openqa.suse.de/tests/8280667#step/image_docker/1932
  https://openqa.suse.de/tests/8280515#step/image_docker/2319
  https://openqa.suse.de/tests/8280373#step/image_docker/771
  https://openqa.suse.de/tests/8280357#step/image_podman/175

• Module Activation via SUSEConnect:
  https://openqa.suse.de/tests/8279393#step/registry/7

Updating system details on https://scc.suse.com ...
Activating PackageHub 15.3 x86_64 ...
Error: Cannot parse response from server

#32 - 2022-03-08 08:20 - jlausuch

Today's failures:

Timeouts in installation jobs:
• https://openqa.suse.de/tests/8287090#step/scc_registration/8 (504 Gateway Time-out)
• https://openqa.suse.de/tests/8288659#step/scc_registration/17 (504 Gateway Time-out)
• https://openqa.suse.de/tests/8286234#step/scc_registration/17 (Cannot parse data from server)

SUSEConnect --list-extensions:
• https://openqa.suse.de/tests/8285389#step/suseconnect_scc/12 (Error: Cannot parse response from server)
• https://openqa.suse.de/tests/8285390#step/suseconnect_scc/12 (Error: Cannot parse response from server)
• https://openqa.suse.de/tests/8285402#step/suseconnect_scc/12 (Error: Cannot parse response from server)
• https://openqa.suse.de/tests/8285445#step/suseconnect_scc/12 (Error: Cannot parse response from server)

Zypper timeouts:
• https://openqa.suse.de/tests/8285493#step/toolbox/87 (toolbox run -c devel -- zypper lr' timed out)

Other connectivity issues:
• https://openqa.suse.de/tests/8285630#step/prepare_instance/45 (https://compute.googleapis.com - network is unreachable)

#33 - 2022-03-08 10:58 - jlausuch

jlausuch wrote:

SUSEConnect --list-extensions:
• https://openqa.suse.de/tests/8285389#step/suseconnect_scc/12 (Error: Cannot parse response from server)
• https://openqa.suse.de/tests/8285390#step/suseconnect_scc/12 (Error: Cannot parse response from server)
• https://openqa.suse.de/tests/8285402#step/suseconnect_scc/12 (Error: Cannot parse response from server)
• https://openqa.suse.de/tests/8285445#step/suseconnect_scc/12 (Error: Cannot parse response from server)

This PR will help with this specific issue:
https://github.com/os-autoinst/os-autoinst-distri-opensuse/pull/14437

#35 - 2022-03-10 07:07 - jlausuch

2 more parent installation jobs today:
• https://openqa.suse.de/tests/8305172#step/scc_registration/20 (504 Gateway Time-out)
• https://openqa.suse.de/tests/8305850#step/scc_registration/16 (Cannot parse data from server)

#36 - 2022-03-15 07:00 - jlausuch

https://openqa.suse.de/tests/8327428#step/scc_registration/33 (Cannot parse data from server)
https://openqa.suse.de/tests/8326890#step/scc_registration/53 (Refreshing service 'Web_and_Scripting_Module_12_x86_64' failed)

SCC issues from today:

- https://openqa.suse.de/tests/8338397#step/scc_registration/7 (Timeout exceeded accessing scc.suse.com)
- https://openqa.suse.de/tests/8336864#step/scc_registration/16 (Cannot parse data from server)

Non-SCC related:

- https://openqa.suse.de/tests/8336767#step/curl_ipv6/4 (could not resolve host www.zq1.de)
- https://openqa.suse.de/tests/8336791#step/wget_ipv6/9 (wget -O -q www3.zq1.de/test.txt' timed out)

I propose to use Nagios to monitor availability of mentioned services (if it is not used already and I just don't know about it) for a week or two to collect good statistics. This topic seems also related to https://progress.opensuse.org/issues/103656 #103656

Yeah, why not, continuous monitoring can give us a hint of those random hiccups.

I have prepared PR to workaround the SCC error pop-ups e.g. https://openqa.suse.de/tests/8368303#step/scc_registration/11
Question is if it's safe to just press OK and continue so I created https://bugzilla.suse.com/show_bug.cgi?id=1197380

We need statistics for what ? To confirm there is network/SCC issue ?

To get some statistics you can look into the data that is available within openQA. We also have https://maintenance-statistics.dyn.cloud.suse.de/question/319 which parses the openQA database.

The graph shows that by far "scc_registration" is now the most failing test module. That was certainly different some months ago.
Sorry, the assignee was removed as I added comment from unupdated session. Great we have statistics proving serious network/SCC issue, time to investigate/fix it.

Since SCC is not the only service failing randomly with timeout, I would like to compare results of SCC monitoring with availability from OpenQA side. That would clearly show whether the problem is in connection of OpenQA network to outside. Might it be correlated with recent moving of servers in Nurnberg from one floor to another? Maybe we will find a long CAT4 cable with many loops around a switching power source or some other computer black magic. In attachment you can see an expert who could help :)

Since SCC is not the only service failing randomly with timeout, I would like to compare results of SCC monitoring with availability from OpenQA side. That would clearly show whether the problem is in connection of OpenQA network to outside.

Yes, we will look into a bit more monitoring on that side. We should be aware that "from OpenQA side" should be distinguished from which machine specially we try to reach scc.suse.com and any CDN component even outside SUSE. I am thinking of connectivity monitoring from each OpenQA worker machine to scc.suse.com

Might it be correlated with recent moving of servers in Nurnberg from one floor to another? […]

This seems to be more unlikely. For example there is https://openqa.suse.de/tests/8368303#step/scc_registration/11 running on openqaworker13, openqaworker13 and most productive openQA workers are located in SUSE Nbg SRV1. The move of QA machines from QA labs was to SUSE Nbg SRV2 which should not significantly affect network capabilities in SRV1.

We discussed this ticket in the weekly SUSE QE sync meeting. okurz suggested to use in particular https://github.com/os-autoinst/scripts/blob/master/README.md#auto-review---automatically-detect-known-issues-in-openqa-jobs-label-openqa-jobs-with-ticket-references-and-optionally-retrigger, see #108773 as an example for that

EDIT:

- https://openqa.suse.de/tests/8379417#step/journal_check/4 is a test on "svirt-xen-hvm" which is running on openqaw5-xen.qa.suse.de so definitely QA network is involved. If DHCP fails that might be openqaw5-xen.qa.suse.de (one could take a look into logs there) but for DNS resolution problem trying to resolve scc.suse.com for sure the DNS server on qanet.qa.suse.de is involved. So we could take a look into DNS server logs on qanet.qa.suse.de (named service).
- https://openqa.suse.de/tests/8380623#step/podman/180 is a failure in podman within an internal tumbleweed container failing to get valid data from download.opensuse.org from within a qemu VM running on openqaworker8 so neither scc.suse.com nor the SUSE QA net are involved.

We could separate into three different areas and follow up with improvements and better investigation for each of these:

- Check DNS resolution and DHCP stability within QA net, e.g. more monitoring, mtr, etc. -> #108845
- Check accessibility to SCC: That should be covered by #102945 but maybe we can progress faster than waiting for "management level" followup with implementing some monitoring on our side, e.g. telegraf ping checks to components like download.opensuse.org, scc.suse.com, proxy-scc

Check DNS resolution and DHCP stability within QA net, e.g. more monitoring, mtr, etc. -> #108845
Check accessibility to SCC: That should be covered by #102945 but maybe we can progress faster than waiting for "management level" followup with implementing some monitoring on our side, e.g. telegraf ping checks to components like download.opensuse.org, scc.suse.com, proxy-scc

HA MM jobs are affected by this since last build 113.1 as well. What I am seeing are SCC connection timeouts: https://openqa.suse.de/tests/8373145#step/welcome/10

But as well hostname resolution problems not only related to SCC, but as well connection between scsi client/server: https://openqa.suse.de/tests/8373191#step/iscsi_client/13

2022-09-04
#52 - 2022-03-24 07:12 - jlausuch
And a non-scc related:
https://openqa.suse.de/tests/8380471#step/rootless_podman/103
Timeout exceeded when accessing 'http://download.opensuse.org/tumbleweed/repo/non-oss/content'
and also https://openqa.suse.de/tests/8380623#step/podman/190

#53 - 2022-03-24 07:42 - jlausuch
This is definitely something related to our infra network: https://openqa.suse.de/tests/8379417#step/journal_check/4
localhost wicked[1059]: eth0: DHCP4 discovery failed
and the next module fails at adding a repo, even after a few retries: https://openqa.suse.de/tests/8379417#step/pam/51
and also, the curl command to upload the logs to the worker: https://openqa.suse.de/tests/8379417#step/pam/54

And DNS resolution issue:
https://openqa.suse.de/tests/8379451#step/suseconnect_scc/16 (getaddrinfo: Temporary failure in name resolution)

#54 - 2022-03-24 09:27 - jlausuch
Today is just a mess:
https://openqa.suse.de/tests/8379452#step/suseconnect_scc/16
https://openqa.suse.de/tests/8379449#step/suseconnect_scc/16
https://openqa.suse.de/tests/8381692#step/docker_runc/116
https://openqa.suse.de/tests/8382172#step/docker/89
https://openqa.suse.de/tests/8382019#step/buildah/78
https://openqa.suse.de/tests/8381902#step/zypper_docker/96
https://openqa.suse.de/tests/8379451#step/patch_and_reboot/148
https://openqa.suse.de/tests/8381666#step/zypper_docker/96

and this is just a portion of all the failures that I've seen today...

#55 - 2022-03-24 09:47 - lpalovsky
Latest build 116.4 is again very bad for HA. Issues are mostly related to either lost connection or DNS resolution failure. This happens between cluster nodes, ssh connection to SUT or to external site like openqa.suse.de or scc.

Various SSH disconnections:
https://openqa.suse.de/tests/8375657#step/patch_sle/150
https://openqa.suse.de/tests/8374846#step/boot_to_desktop/16

Node being unreachable/network resolution issue:
https://openqa.suse.de/tests/8376302#step/ha_cluster_init/15
https://openqa.suse.de/tests/8376293#step/iscsi_client/11
https://openqa.suse.de/tests/8363004#step/ha_cluster_init/16

Examples above are only a portion of failed tests.

#56 - 2022-03-24 10:29 - okurz
So in the above examples I can see

- qemu jobs on machines within SRV1 having problems to access dist.suse.de, e.g. https://openqa.suse.de/tests/8381692#step/docker_runc/116 (aarch64) and https://openqa.suse.de/tests/8382172#step/docker/91 (x86_64) -> please create a specific subticket about that and investigate the network performance between the physical machines and dist.suse.de or create EngInfra ticket
- XEN-svirt jobs on openqaw5-xen qa.suse.de, e.g. https://openqa.suse.de/tests/8379451#step/patch_and_reboot/148, that is covered in #108845 , feel welcome to add an according auto-review regex for the error messages. We are looking into the problem
- s9x0 jobs like https://openqa.suse.de/tests/8375657#step/patch_sle/150 loosing connections just in the middle -> please create a specific subticket about that and investigate the network performance between the physical machines and dist.suse.de or create EngInfra ticket. Could be related to #99345 although the mentioned job is failed, not incomplete with "Connection timed out"
- ppc qemu jobs like https://openqa.suse.de/tests/8376302#step/ha_cluster_init/15 -> that again looks to be something we could treat as separate. Please report a ticket as such.

#57 - 2022-03-24 18:42 - okurz
- Related to action #108668: Failed systemd services alert (except openqa.suse.de) for < 60 min added

#59 - 2022-03-25 07:13 - jlausuch
More name resolution ones:
https://openqa.suse.de/tests/8367036#step/suseconnect_scc/16 - openqaworker2:9
And DHCP discovery failures:
https://openqa.suse.de/tests/8387055#step/journal_check/4 - openqaworker2:16
https://openqa.suse.de/tests/8387085#step/journal_check/4 - openqaworker2:16

#60 - 2022-03-25 07:53 - lpalovsky
okurz wrote:

- ppc qemu jobs like https://openqa.suse.de/tests/8376302#step/ha_cluster_init/15 -> that again looks to be something we could treat as separate. Please report a ticket as such.

Thanks, I have created a separate ticket for that issue: #108962

#61 - 2022-03-28 09:48 - okurz
A change in network cabling was made around 2022-03-28 0815Z. We now observed a stable connection between the QA switches and previously affected machines. We will continue to monitor the situation. See #108845#note-25 for details. Please report any related issues still occuring for openQA tests that have been started after the above time.

#62 - 2022-03-29 07:35 - jlausuch
I noticed less network issues lately, but still some DNS resolution failures here and there (not so many):

- https://openqa.suse.de/tests/8426618#step/cifs/63 could not resolve address for currywurst.qam.suse.de
- https://openqa.suse.de/tests/8426620#step/podman/96 could not resolve host google.de

#63 - 2022-04-06 11:47 - jstehlik
I consider this issue resolved by fixing cabling. The topic of infrastructure reliability shall be further followed in https://progress.opensuse.org/issues/109250.

#64 - 2022-04-06 11:50 - jstehlik
- Status changed from Workable to Feedback

#65 - 2022-04-09 17:04 - okurz
jstehlik wrote:

I consider this issue resolved by fixing cabling. The topic of infrastructure reliability shall be further followed in https://progress.opensuse.org/issues/109250.

I think we shouldn't just declare this ticket as "resolved" just because one problem was fixed. The problem might have been small but the impact was huge. I suggest we discuss further improvement ideas on multiple levels so that any new or next upcoming problem would have less severe impact. Also multiple subtasks are sill open. If you like to await their results before reviewing this ticket then I suggest to use the "Blocked" status.

#66 - 2022-04-13 09:09 - dzedro
- File Screenshot_2022-04-12_22-16-57.png added

Now there are serious network problems on s390x, this kind of issue happened also before, but now it's very bad, I guess since yesterday afternoon. If I look on the live log or serial output of running job, at some point it will freeze no boot update or worker debug output. Then just fail will show up, see attachment.

Three variations of test unable to boot.
https://openqa.suse.de/tests/8538657#step/bootloader_zkvm/18
https://openqa.suse.de/tests/8538655#step/boot_to_desktop/16
https://openqa.suse.de/tests/8538654#step/bootloader_start/19

#67 - 2022-04-14 08:06 - dzedro
Also today there are many failed s390x tests at boot e.g.
https://openqa.suse.de/tests/8550263#step/bootloader_start/19
https://openqa.suse.de/tests/8550253#step/bootloader_zkvm/18

#68 - 2022-04-14 08:41 - dzedro
Looks like the s390x failures were not related to network but disk space. https://suse.slack.com/archives/C02CANHLANP/p1649923742408719
We discussed within the weekly QE Sync meeting 2022-05-04 that some problems have been fixed (operational work). We should raise it to LSG mgmt, e.g. in the "Open Doors meeting" how we suffer from the impact of bad infrastructure and that we motivate for improvements, e.g. with the planned datacenter move. szarate mentions that we already occupying multiple engineers in their daily work to investigate network related problems, e.g. effectively loosing a full day due to a non-executed milestone build validation.

EDIT: I got confirmation from jpupava and jlausuch that both see the original issue(s) resolved. We agree that there is still a lot of room for improvement, e.g. base-level infrastructure monitoring but that has been stated likely sufficiently. I think it makes sense to track the work in the still open subtasks before resolving the ticket but after all subtasks are resolved then nothing more would be needed. Added acceptance criteria as such.

EDIT: I got confirmation from jpupava and jlausuch that both see the original issue(s) resolved. We agree that there is still a lot of room for improvement, e.g. base-level infrastructure monitoring but that has been stated likely sufficiently. I think it makes sense to track the work in the still open subtasks before resolving the ticket but after all subtasks are resolved then nothing more would be needed. Added acceptance criteria as such.

On the topic of how much in terms of hours this took:

While we know what the root cause of the problem was (The network cable), there's a second level of the root cause which is lack of manpower and monitoring on the infrastructure (I guess here it's SUSE IT and QE).

Now in terms of costs:

So, It's around 528 hours, only counting 3 engineers from QE tools, 3 from openQA maintenance review, accounting for ~59 QE Engineers, 1 RM, 66 persons in total, assuming that:

- Some of them had to do twice review (Day of the failure and day after).
- Some of them were sitting idly during manual validation, due to some resources not being available.
- Some of them could not verify or run verification runs that depended on resources in OSD or qanet due to the network being degraded/down.

This doesn't account for the amount of hours of automated testing, that can be calculated, and directly converted to €, if we take into account the power consumption of the servers, nor the work that had to be done as fall out (Your time, meeting last wed, + any other meeting happening).

NOTE: My numbers might be off, if somebody wants to cross check, be my guest :). Also having the data of how many hours of testing on openQA were lost, would be kind of a good idea, for this and further incidents.

sزارate from today's discussion it sounds like some people still see some issues, you mentioned "zypper problems". I think the current subtasks to this ticket do not cover this. We need to ensure that there are tickets which handle each issue explicitly so that no one is waiting for miracle solution problems.

There is further evaluation needed when new test rack is set up in Prague.

Also this issue seems related: QA network infrastructure Ping time alert for several hosts https://progress.opensuse.org/issues/113498

Related to action #113528: [qe-core] test fails in bootloader_zkvm - performance degradation in the s390 network is causing serial console to be unreliable (and killing jobs slowly) added

Related to action #113716: [qe-core] proxy-scc is down added