

openQA Infrastructure - action #90968

[alert] Multiple flaky incomplete job alerts on Sunday

2021-04-12 07:54 - cdywan

Status:	Resolved	Start date:	2021-04-12
Priority:	High	Due date:	2021-04-29
Assignee:	mkittler	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:	Ready		
Description			
Incomplete jobs (not restarted) of last 24h alert - Ok after 2 minutes			
Metric name			
Value			
Queue: State (SUSE) alert* - OK after 3 minutes			
Error message			
tsdb.HandleRequest() error Get "http://localhost:8086/query?db=telegraf&epoch=s&q=SELECT+mean%28%22scheduled%22%29+FROM+%22openqa_jobs%22+WHERE+%22url%22+%3D+%27https%3A%2F%2Fopenqa.suse.de%27+AND+time+%3E+now%28%29+--+1m+GROUP+BY+time%2840s%29+fill%28null%29": dial tcp [::1]:8086: connect: connection refused			
Metric name			
Value			
New incompletes alert - OK after 3 minutes			
Error message			
tsdb.HandleRequest() error Get "http://localhost:8086/query?db=telegraf&epoch=s&q=SELECT+non_negative_difference%28distinct%28%22incompletes_last_24h%22%29%29+FROM+%22postgresql%22+WHERE+time+%3E+now%28%29+--+1m+GROUP+BY+time%2850ms%29": dial tcp [::1]:8086: connect: connection refused			
Metric name			
Value			

History

#1 - 2021-04-12 08:03 - okurz

- Priority changed from Normal to High
- Target version set to Ready

#2 - 2021-04-12 08:16 - cdywan

- Description updated

#3 - 2021-04-12 08:36 - okurz

- Subject changed from Multiple flaky incomplete job alerts on Sunday to [alert] Multiple flaky incomplete job alerts on Sunday
- Status changed from New to Workable

#4 - 2021-04-12 13:32 - mkittler

This was caused by InfluxDB being shortly restarting and not by a high amount of incompletes:

```
martchus@openqa-monitor:~> systemctl status influxdb.service
● influxdb.service - InfluxDB database server
   Loaded: loaded (/usr/lib/systemd/system/influxdb.service; enabled; vendor preset: disabled)
   Active: active (running) since Sun 2021-04-11 03:30:36 CEST; 1 day 11h ago
   Main PID: 1668 (influxd)
```

```
Tasks: 12
CGroup: /system.slice/influxdb.service
└─1668 /usr/bin/influxd -config /etc/influxdb/config.toml -pidfile /run/influxdb/influxdb.pid
```

The whole system was actually restarting:

```
martchus@openqa-monitor:~> sudo journalctl --system --boot
-- Logs begin at Wed 2021-04-07 18:19:11 CEST, end at Mon 2021-04-12 15:30:06 CEST. --
Apr 11 03:30:20 openqa-monitor kernel: Linux version 5.3.18-lp152.69-default (geeko@buildhost) (gcc version 7.5.0 (SUSE Linux)) #1 SMP Tue Apr 6 11:41:13 UTC 2021 (d532e33)
```

So it looks like InfluxDB wasn't just ready soon enough. Maybe we can just increase the grace period for InfluxDB being inaccessible.

#5 - 2021-04-14 09:35 - mkittler

- Status changed from Workable to In Progress

- Assignee set to mkittler

Maybe we can just increase the grace period for InfluxDB being inaccessible.

There's not really a grace period for that but I'll try to set the "No Data & Error Handling" for "If execution error or timeout" to "Keep last state". If InfluxDB is broken longer this should be caught by the failed systemd services alert anyways. I'll try to test whether it actually includes the monitoring host by intentionally failing a unit.

#6 - 2021-04-14 10:02 - okurz

as discussed an alternative might be to overwrite the grafana with influxdb-probing, waiting until that is done and only then executing the expected service as so far we seem to have seen this problem only when influxdb+grafana starts up, not anytime during normal execution

#7 - 2021-04-15 04:13 - openqa_review

- Due date set to 2021-04-29

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

#8 - 2021-04-15 14:19 - mkittler

SR: https://gitlab.suse.de/openqa/salt-states-openqa/-/merge_requests/479

#9 - 2021-04-15 14:54 - okurz

I will try https://gitlab.suse.de/openqa/salt-states-openqa/-/merge_requests/477 now. be ready for any alarm explosion

<https://stats.openqa-monitor.qa.suse.de/d/nRDab3Jiz/openqa-jobs-test?orgId=1&editPanel=17&tab=alert> states it has no data since 2 minutes, maybe telegraf config is now borked.

Marius Kittler: Apr 15 16:32:26 openqa telegraf[17485]: 2021-04-15T14:32:26Z W! Telegraf is not permitted to read /etc/telegraf/telegraf.d
The dir is there. I've just been restarting telegraf and now it seems to work. Maybe salt restarted telegraf and only then created the directory.

Oliver Kurz: I suspected the same. I just triggered a restart of all telegraf services on all salt nodes as well. I can add a restart trigger in salt easily as well

Manually executing

```
telegraf --debug -config /etc/telegraf/telegraf.conf -config-directory /etc/telegraf/telegraf.d --test 2>&1 |
grep '\(http_response\|systemd_failed\)'
```

looks good as well.

After some minutes it looks good again.

#10 - 2021-04-16 08:37 - mkittler

- Status changed from In Progress to Feedback

My SR has been merged as well.

#11 - 2021-04-19 09:48 - okurz

- Status changed from Feedback to Resolved

crosschecked. The new dashboard <https://monitor.qa.suse.de/d/EML0bpuGk/monitoring?orgId=1> is there and maintained in salt. All good then.