openQA Infrastructure - action #66019

Configure downsampling and a retention policy for InfluxDB

2020-04-23 13:40 - mkittler

Status:	Workable		**Start date:**	2020-04-23
Priority:	Low		**Due date:**	
Assignee:			**% Done:**	0%
Category:			**Estimated time:**	0.00 hour
Target version:	future			

Description

Motivation

When working on #64574 I've noticed that we don't have a retention policy and the default policy seems to be storing the data forever. That's likely something we want to adjust in the long term.

Acceptance criteria

- **AC1:** Downsampling and retention is enabled with at least *some* sensible settings on the influxdb on monitor.qa.suse.de

Suggestion

- Read https://docs.influxdata.com/influxdb/v2.0/process-data/common-tasks/downsample-data/ about downsampling
- Read https://docs.influxdata.com/influxdb/v2.0/reference/api/influxdb-1x/dbrp/ about retention
- Look into what mkittler has drafted so far (not tested): https://gitlab.suse.de/mkittler/salt-states-openqa/-/commit/1880714e73d467ac5ed9c0a5b5cec4467fc14a20
- Introduce the corresponding changes to limit the storage space used by influxdb, mainly to prevent the storage to run full

History

#1 - 2020-07-29 07:13 - okurz

- Priority changed from Normal to Low

#2 - 2020-10-22 13:36 - okurz

- Target version set to Ready

#3 - 2020-11-05 15:03 - okurz

- Tags set to monitor, influxdb, infrastructure, administration, osd

- Subject changed from Configure a retention policy for InfluxDB to Configure downsampling and a retention policy for InfluxDB

- Description updated

- Status changed from New to Workable

#4 - 2020-11-05 15:12 - okurz

found https://towardsdatascience.com/influxdb-data-retention-f026496d708f which is describing a bit of an overview. I have not really understood if we need to define something in a config file or set in the database itself, if we need to run some additional process on a regular base or if that is handled within influxdb as well.

#5 - 2020-11-27 19:07 - okurz

- Target version changed from Ready to future

```
$ ssh monitor.qa df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/vda2        77G   27G   50G  35% /
/dev/vdb1       300G   87G  212G  30% /var/lib/influxdb
```

I think we have enough headroom for now. Removing from backlog again due to backlog capacity constraints.