

openQA Infrastructure - action #103524

OW1: performance loss size:M

2021-12-06 08:33 - dimstar

Status: Resolved	Start date: 2021-12-06
Priority: High	Due date:
Assignee: mkittler	% Done: 100%
Category:	Estimated time: 0.00 hour
Target version: Ready	

Description

Observation

A lot of tests are failing since Friday, even the rather 'simple' raid install tests. The failure happens very often in 'await_install', after having spent 1 hour being busy installing RPMs. Those tests generally pass in 20 minutes through the installer.

This started happening across the product to be tested as well as also in stagings, at the same time, which makes 'product-related issues' less probable than ow1 related issues.

Acceptance criteria

- **AC1:** openqaworker1 is back in production with a stable set of worker instances

Suggestions

- Reduce the number of openQA worker instances until we don't commonly have failing or incompleting jobs
- If there are recurring I/O errors showing up in logs consider replacing hardware again or the other NVMe if that is faulty

Rollback actions

- Use a proper RAID again of all NVMe devices for /var/lib/openqa (fvogdt has used an unusual direct use of /dev/nvme0n1 for /var/lib/openqa for now)
- Enable additional worker instances again after hardware replacement: `systemctl unmask --now openqa-worker-auto-restart@{16..20}`
- Increase WORKERCACHE in openqaworker1:/etc/openqa/workers.ini to a higher value, e.g. 400GB again after replacement of NVMe an increase of space
- Increase number of worker instances again, i.e. enable worker instance [7..16]

Related issues:

Related to openQA Project - action #103581: Many jobs on openqa.opensuse.org ...	Resolved	2021-12-07
Has duplicate openQA Infrastructure - action #107017: Random asset download (...)	Resolved	
Copied to openQA Infrastructure - action #107017: Random asset download (cach...	Resolved	

History

#1 - 2021-12-06 09:10 - favogt

There are disk issues:

```
openqaworker1:/etc # smartctl -a /dev/nvme0n1
smartctl 7.0 2019-05-21 r4917 [x86_64-linux-5.3.18-59.34-default] (SUSE RPM)
Copyright (C) 2002-18, Bruce Allen, Christian Franke, www.smartmontools.org
```

```
=== START OF INFORMATION SECTION ===
```

```
Model Number:          INTEL SSDPE2ME400G4
Serial Number:         CVMD4352001W400FGN
Firmware Version:      8DV10151
PCI Vendor/Subsystem ID: 0x8086
IEEE OUI Identifier:   0x5cd2e4
Controller ID:         0
Number of Namespaces:  1
Namespace 1 Size/Capacity: 400.088.457.216 [400 GB]
Namespace 1 Formatted LBA Size: 4096
Local Time is:         Mon Dec 6 10:02:32 2021 CET
```

Firmware Updates (0x02): 1 Slot
Optional Admin Commands (0x0006): Format Frmw_DL
Optional NVM Commands (0x0006): Wr_Unc DS_Mngmt
Maximum Data Transfer Size: 32 Pages

Supported Power States

St Op	Max	Active	Idle	RL	RT	WL	WT	Ent_Lat	Ex_Lat
0 +	25.00W	-	-	0	0	0	0	0	0

Supported LBA Sizes (NSID 0x1)

Id Fmt	Data	Metadt	Rel_Perf
0 -	512	0	2
1 -	512	8	2
2 -	512	16	2
3 +	4096	0	0
4 -	4096	8	0
5 -	4096	64	0
6 -	4096	128	0

=== START OF SMART DATA SECTION ===

SMART overall-health self-assessment test result: PASSED

SMART/Health Information (NVMe Log 0x02)

Critical Warning: 0x00
Temperature: 24 Celsius
Available Spare: 99%
Available Spare Threshold: 10%
Percentage Used: 32%
Data Units Read: 1.123.143.620 [575 TB]
Data Units Written: 2.491.413.126 [1,27 PB]
Host Read Commands: 7.436.632.694
Host Write Commands: 12.539.489.915
Controller Busy Time: 10.331
Power Cycles: 101
Power On Hours: 55.231
Unsafe Shutdowns: 21
Media and Data Integrity Errors: 0
Error Information Log Entries: 0

Error Information (NVMe Log 0x01, max 64 entries)

Num	ErrCount	SQId	CmdId	Status	PELoc	LBA	NSID	VS
0	1	13	0x0000	0x4502	0x000	0	1	-

openqaworker1:/etc # smartctl -a /dev/nvme1n1
smartctl 7.0 2019-05-21 r4917 [x86_64-linux-5.3.18-59.34-default] (SUSE RPM)
Copyright (C) 2002-18, Bruce Allen, Christian Franke, www.smartmontools.org

=== START OF INFORMATION SECTION ===

Model Number: INTEL SSDPEKNW010T8
Serial Number: BTNH022600SN1POB
Firmware Version: 002C
PCI Vendor/Subsystem ID: 0x8086
IEEE OUI Identifier: 0x5cd2e4
Controller ID: 1
Number of Namespaces: 1
Namespace 1 Size/Capacity: 1.024.209.543.168 [1,02 TB]
Namespace 1 Formatted LBA Size: 512
Local Time is: Mon Dec 6 10:02:34 2021 CET
Firmware Updates (0x14): 2 Slots, no Reset required
Optional Admin Commands (0x0017): Security Format Frmw_DL Self_Test
Optional NVM Commands (0x005f): Comp Wr_Unc DS_Mngmt Wr_Zero Sav/Sel_Feat Timestmp
Maximum Data Transfer Size: 32 Pages
Warning Comp. Temp. Threshold: 77 Celsius
Critical Comp. Temp. Threshold: 80 Celsius

Supported Power States

St Op	Max	Active	Idle	RL	RT	WL	WT	Ent_Lat	Ex_Lat
0 +	4.00W	-	-	0	0	0	0	0	0
1 +	3.00W	-	-	1	1	1	1	0	0
2 +	2.20W	-	-	2	2	2	2	0	0
3 -	0.0300W	-	-	3	3	3	3	5000	5000
4 -	0.0040W	-	-	4	4	4	4	5000	9000

Supported LBA Sizes (NSID 0x1)

Id Fmt	Data	Metadt	Rel_Perf
--------	------	--------	----------

```
0 +      512      0      0
```

```
=== START OF SMART DATA SECTION ===
```

```
SMART overall-health self-assessment test result: FAILED!  
- NVM subsystem reliability has been degraded
```

```
SMART/Health Information (NVMe Log 0x02)
```

```
Critical Warning:      0x04  
Temperature:          32 Celsius  
Available Spare:      100%  
Available Spare Threshold: 10%  
Percentage Used:      118%  
Data Units Read:      960.439.256 [491 TB]  
Data Units Written:   1.534.101.586 [785 TB]  
Host Read Commands:   6.131.723.556  
Host Write Commands:  6.981.165.126  
Controller Busy Time: 322.477  
Power Cycles:         5  
Power On Hours:       10.708  
Unsafe Shutdowns:    4  
Media and Data Integrity Errors: 0  
Error Information Log Entries: 0  
Warning Comp. Temperature Time: 0  
Critical Comp. Temperature Time: 0
```

```
Error Information (NVMe Log 0x01, max 256 entries)
```

```
No Errors Logged
```

```
openqaworker1:/etc # cat /proc/mdstat
```

```
Personalities : [raid1] [raid0]  
md0 : active raid0 sdb2[0] sda2[1]  
      2115520 blocks super 1.0 32k chunks
```

```
md127 : active raid0 nvme0n1[0] nvme1n1[1]  
      1390651392 blocks super 1.2 512k chunks
```

```
md1 : active raid1 sdb3[0] sda3[1]  
      975701824 blocks super 1.0 [2/2] [UU]  
      bitmap: 0/8 pages [0KB], 65536KB chunk
```

```
unused devices: <none>
```

```
openqaworker1:/etc # dmesg | grep md127
```

```
[ 27.406823] md127: detected capacity change from 0 to 1424027025408  
[ 29.791083] EXT4-fs (md127): mounting ext2 file system using the ext4 subsystem  
[ 29.793829] EXT4-fs (md127): warning: mounting unchecked fs, running e2fsck is recommended  
[ 29.800001] EXT4-fs (md127): mounted filesystem without journal. Opts: (null)  
[ 92.009666] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3516, block bitmap and bg descriptor inconsistent: 17811 vs 17814 free clusters  
[ 92.719845] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 4226, block bitmap and bg descriptor inconsistent: 25284 vs 25287 free clusters  
[ 110.632664] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3518, block bitmap and bg descriptor inconsistent: 26031 vs 25834 free clusters  
[ 332.995178] EXT4-fs (md127): error count since last fsck: 2825  
[ 332.995181] EXT4-fs (md127): initial error at time 1629432432: mb_free_blocks:1457: inode 28803076: block 1 15238455  
[ 332.995184] EXT4-fs (md127): last error at time 1638758894: ext4_mb_generate_buddy:747  
[ 4355.696038] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 4241, block bitmap and bg descriptor inconsistent: 23039 vs 23042 free clusters  
[ 4545.295790] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3194, block bitmap and bg descriptor inconsistent: 25197 vs 25217 free clusters  
[ 4607.626626] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3236, block bitmap and bg descriptor inconsistent: 27817 vs 27912 free clusters  
[ 5010.953172] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3485, block bitmap and bg descriptor inconsistent: 16629 vs 16641 free clusters  
[ 5022.249988] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3537, block bitmap and bg descriptor inconsistent: 13818 vs 13820 free clusters  
[ 5439.212789] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3833, block bitmap and bg descriptor inconsistent: 28007 vs 28005 free clusters  
[ 5835.279080] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 4141, block bitmap and bg descriptor inconsistent: 32254 vs 32287 free clusters  
[ 7515.865532] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 5267, block bitmap and bg descriptor inconsistent: 12442 vs 12454 free clusters  
[11877.153446] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 7195, block bitmap and bg descriptor inconsistent: 5553 vs 5666 free clusters  
[11877.191196] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 7196, block bitmap and bg descriptor inconsistent: 5553 vs 5666 free clusters
```

```

iptor inconsistent: 16234 vs 18561 free clusters
[11877.225669] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 7197, block bitmap and bg descr
iptor inconsistent: 18454 vs 18695 free clusters
openqaworker1:/etc # dmesg | grep md1
[ 12.819165] md/raid1:md1: active with 2 out of 2 mirrors
[ 12.852649] md1: detected capacity change from 0 to 999118667776
[ 13.798414] BTRFS: device fsid ff1922d2-d2e4-4860-9634-acac681dd0f9 devid 1 transid 2287684 /dev/md1
[ 13.929720] BTRFS info (device md1): disk space caching is enabled
[ 13.929721] BTRFS info (device md1): has skinny extents
[ 14.475616] BTRFS info (device md1): disk space caching is enabled
[ 18.952912] BTRFS info (device md1): disk space caching is enabled
[ 27.406823] md127: detected capacity change from 0 to 1424027025408
[ 29.520197] BTRFS info (device md1): disk space caching is enabled
[ 29.791083] EXT4-fs (md127): mounting ext2 file system using the ext4 subsystem
[ 29.793829] EXT4-fs (md127): warning: mounting unchecked fs, running e2fsck is recommended
[ 29.800001] EXT4-fs (md127): mounted filesystem without journal. Opts: (null)
[ 92.009666] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3516, block bitmap and bg descr
iptor inconsistent: 17811 vs 17814 free clusters
[ 92.719845] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 4226, block bitmap and bg descr
iptor inconsistent: 25284 vs 25287 free clusters
[ 110.632664] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3518, block bitmap and bg descr
iptor inconsistent: 26031 vs 25834 free clusters
[ 332.995178] EXT4-fs (md127): error count since last fsck: 2825
[ 332.995181] EXT4-fs (md127): initial error at time 1629432432: mb_free_blocks:1457: inode 28803076: block 1
15238455
[ 332.995184] EXT4-fs (md127): last error at time 1638758894: ext4_mb_generate_buddy:747
[ 4355.696038] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 4241, block bitmap and bg descr
iptor inconsistent: 23039 vs 23042 free clusters
[ 4545.295790] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3194, block bitmap and bg descr
iptor inconsistent: 25197 vs 25217 free clusters
[ 4607.626626] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3236, block bitmap and bg descr
iptor inconsistent: 27817 vs 27912 free clusters
[ 5010.953172] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3485, block bitmap and bg descr
iptor inconsistent: 16629 vs 16641 free clusters
[ 5022.249988] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3537, block bitmap and bg descr
iptor inconsistent: 13818 vs 13820 free clusters
[ 5439.212789] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 3833, block bitmap and bg descr
iptor inconsistent: 28007 vs 28005 free clusters
[ 5835.279080] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 4141, block bitmap and bg descr
iptor inconsistent: 32254 vs 32287 free clusters
[ 7515.865532] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 5267, block bitmap and bg descr
iptor inconsistent: 12442 vs 12454 free clusters
[11877.153446] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 7195, block bitmap and bg descr
iptor inconsistent: 5553 vs 5666 free clusters
[11877.191196] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 7196, block bitmap and bg descr
iptor inconsistent: 16234 vs 18561 free clusters
[11877.225669] EXT4-fs error (device md127): ext4_mb_generate_buddy:747: group 7197, block bitmap and bg descr
iptor inconsistent: 18454 vs 18695 free clusters

```

While QEMU was running, the system load exceeded 300 and sync never returned. After stopping QEMU, this situation got much better.

Simply recreating the ext2 as ext4 should help against the errors, but the IO issues are probably caused by the block device itself.

#2 - 2021-12-06 09:39 - favogt

Not the right ticket for this, but it's close enough:

I recovered openqaworker4 from a btrfs error on /. The error caused / to mount only read-only and every boot ended up in the emergency shell. btrfs check /dev/md1 reported errors:

```

Opening filesystem to check...
Checking filesystem on /dev/md1
UUID: 9058d83e-9f1f-49be-a4c0-8f975248ad39
[1/7] checking root items
[2/7] checking extents
ref mismatch on [1054179328 16384] extent item 0, found 1
tree backref 1054179328 parent 15764 root 15764 not found in extent tree
backpointer mismatch on [1054179328 16384]
ref mismatch on [1078231040 16384] extent item 0, found 1
tree backref 1078231040 parent 15764 root 15764 not found in extent tree
backpointer mismatch on [1078231040 16384]
ref mismatch on [1090617344 16384] extent item 0, found 1
tree backref 1090617344 parent 15764 root 15764 not found in extent tree
backpointer mismatch on [1090617344 16384]
ERROR: errors found in extent allocation tree or chunk allocation

```

```

[3/7] checking free space cache
[4/7] checking fs roots
[5/7] checking only csums items (without verifying data)
[6/7] checking root refs
[7/7] checking quota groups
ERROR: out of memory
ERROR: Loading qgroups from disk: -2
ERROR: failed to check quota groups
found 22590820352 bytes used, error(s) found
total csum bytes: 14809236
total tree bytes: 233046016
total fs tree bytes: 198819840
total extent tree bytes: 16171008
btree space waste bytes: 44986480
file data blocks allocated: 1505488662528
referenced 28240560128
extent buffer leak: start 249446400 len 16384
extent buffer leak: start 257523712 len 16384

```

To repair that, it's necessary to have /dev/md1 not mounted. For that, I first tried to reboot and get into the GRUB menu but that was somehow disabled, both on the serial console (IPMI) as well as the graphical one (java console). So I had to use kexec in the emergency shell:

```
kexec --initrd=/boot/initrd --append="console=tty0 console=ttyS1,115200n rd.break" /boot/vmlinuz
```

Inside, I mounted /dev/md127 to store a btrfs metadata backup:

```

sh-4.4# mdadm --assemble --scan
[ 174.399341] md: md0 stopped.
[ 174.409997] md0: detected capacity change from 0 to 2153709568
mdadm: /dev/md/0 has been started with 2 drives.
[ 174.594991] md: md127 stopped.
[ 174.599073] nvme1n1: p1
[ 174.604981] md127: detected capacity change from 0 to 799906201600
mdadm: /dev/md/o[ 174.612954] nvme1n1: p1
penqa has been started with 2 drives.
sh-4.4#
sh-4.4#
sh-4.4# cat /proc/mdstat
Personalities : [raid1] [raid0]
md127 : active raid0 nvme1n1[0] nvme0n1[1]
      781158400 blocks super 1.2 512k chunks

md0 : active raid0 sda1[0] sdb1[1]
      2103232 blocks super 1.0 32k chunks

md1 : active raid1 sdb2[1] sda2[0]
      975708992 blocks super 1.0 [2/2] [UU]
      bitmap: 0/8 pages [0KB], 65536KB chunk

unused devices: <none>
sh-4.4# mount /dev/md127 /mnt
mount: /mnt: mount point does not exist.
sh-4.4# mkdir /mnt
sh-4.4# mount /dev/md127 /mnt
mount: /mnt: unknown filesystem type 'ext2'.
sh-4.4#
sh-4.4# mount /dev/md1 -o ro /sysroot
[ 210.003355] BTRFS info (device md1): disk space caching is enabled
[ 210.009570] BTRFS info (device md1): has skinny extents
... (omitting some dead-ends and mistypes)
sh-4.4# chroot /sysroot modprobe ext4
sh-4.4# mount /dev/md127 /mnt
sh-4.4# cp /sysroot/usr/sbin/btrfs-image /tmp
sh-4.4# umount /sysroot
sh-4.4# /tmp/btrfs-image -w /dev/md1 /mnt/btrfs-image-2
sh-4.4# ls -lh /mnt/btrfs-image*
-rw-r--r-- 1 root root 1.3G Dec  6 09:00 /mnt/btrfs-image
-rw-r--r-- 1 root root 1.3G Dec  6 09:15 /mnt/btrfs-image-2

```

And finally ran btrfs check --repair:

```

sh-4.4# btrfs check --repair /dev/md1
enabling repair mode
Opening filesystem to check...

```

```

Checking filesystem on /dev/md1
UUID: 9058d83e-9f1f-49be-a4c0-8f975248ad39
[1/7] checking root items
Fixed 0 roots.
[2/7] checking extents
ref mismatch on [1054179328 16384] extent item 0, found 1
tree backref 1054179328 parent 15764 root 15764 not found in extent tree
backpointer mismatch on [1054179328 16384]
adding new tree backref on start 1054179328 len 16384 parent 0 root 15764
Repaired extent references for 1054179328
ref mismatch on [1078231040 16384] extent item 0, found 1
tree backref 1078231040 parent 15764 root 15764 not found in extent tree
backpointer mismatch on [1078231040 16384]
adding new tree backref on start 1078231040 len 16384 parent 0 root 15764
Repaired extent references for 1078231040
ref mismatch on [1090617344 16384] extent item 0, found 1
tree backref 1090617344 parent 15764 root 15764 not found in extent tree
backpointer mismatch on [1090617344 16384]
adding new tree backref on start 1090617344 len 16384 parent 0 root 15764
Repaired extent references for 1090617344
No device size related problem found
[3/7] checking free space cache
cache and super generation don't match, space cache will be invalidated
[4/7] checking fs roots
[5/7] checking only csums items (without verifying data)
[6/7] checking root refs
[7/7] checking quota groups
ERROR: out of memory
ERROR: Loading qgroups from disk: -2
ERROR: failed to check quota groups
found 22590820352 bytes used, error(s) found
total csum bytes: 14809236
total tree bytes: 233046016
total fs tree bytes: 198819840
total extent tree bytes: 16171008
btree space waste bytes: 44986480
file data blocks allocated: 1505488662528
referenced 28240560128
extent buffer leak: start 163807232 len 16384
extent buffer leak: start 257523712 len 16384

```

After this, btrfs check was mostly happy:

```

sh-4.4# btrfs check /dev/md1
Opening filesystem to check...
ERROR: mount check: cannot open /dev/md1: No such file or directory
ERROR: could not check mount status: No such file or directory
sh-4.4# btrfs check /dev/md1
Opening filesystem to check...
Checking filesystem on /dev/md1
UUID: 9058d83e-9f1f-49be-a4c0-8f975248ad39
[1/7] checking root items
[2/7] checking extents
[3/7] checking free space cache
cache and super generation don't match, space cache will be invalidated
[4/7] checking fs roots
[5/7] checking only csums items (without verifying data)
[6/7] checking root refs
[7/7] checking quota groups
ERROR: out of memory
ERROR: Loading qgroups from disk: -2
ERROR: failed to check quota groups

```

Finally I did a qgroup rescan (mount /dev/md1 /sysroot; btrfs quota rescan -w /sysroot and rebooted. The system came up good.

I changed GRUB settings to the default, in the hope that this fixes the missing menu.

#3 - 2021-12-06 10:33 - favogt

- Status changed from New to In Progress

- Assignee set to favogt

- Priority changed from Normal to Urgent

Back to the original topic, openqaworker1.

The NVME health doesn't look good, so what might help is to do a "factory reset" on both, a full erase cycle. For that I entered a transactional-update shell to install nvme-cli and ran:

```
transactional update # mdadm --stop /dev/md127
mdadm: stopped /dev/md127
transactional update # nvme list
Node          SN              Model          Namespace Usage
  Format      FW Rev
-----
/dev/nvme0n1  CVMD4352001W400FGN  INTEL SSDPE2ME400G4  1      400.09 GB / 400.09 GB
  4 KiB + 0 B  8DV10151
/dev/nvme1n1  BTNH022600SN1P0B   INTEL SSDPEKNW010T8  1      1.02 TB / 1.02 TB
  512 B + 0 B  002C
transactional update # nvme id-ctrl -H /dev/nvme0 | grep Format
[1:1] : 0x1  Format NVM Supported
[0:0] : 0    Admin Vendor Specific Commands uses Vendor Specific Format
[0:0] : 0x1  Format Applies to All Namespace(s)
[0:0] : 0    NVM Vendor Specific Commands uses Vendor Specific Format
transactional update # nvme id-ctrl -H /dev/nvme1 | grep Format
[1:1] : 0x1  Format NVM Supported
[0:0] : 0    Admin Vendor Specific Commands uses Vendor Specific Format
[0:0] : 0    Format Applies to Single Namespace(s)
[0:0] : 0    NVM Vendor Specific Commands uses Vendor Specific Format
transactional update # nvme format -r --ses=1 /dev/nvme0
You are about to format nvme0, namespace 0xffffffff(ALL namespaces).
Controller nvme0 has child namespace(s):nvme0n1
```

```
WARNING: Format may irrevocably delete this device's data.
You have 10 seconds to press Ctrl-C to cancel this operation.
```

```
Use the force [--force|-f] option to suppress this warning.
```

```
Sending format operation ...
```

```
Success formatting namespace:ffffffff
```

```
transactional update # nvme format -r --ses=1 /dev/nvme1
```

```
get-namespace-id: Inappropriate ioctl for device
```

```
transactional update # nvme format -r --ses=1 -n 1 /dev/nvme1
```

```
You are about to format nvme1, namespace 0x1.
```

```
Controller nvme1 has child namespace(s):nvme1n1
```

```
WARNING: Format may irrevocably delete this device's data.
You have 10 seconds to press Ctrl-C to cancel this operation.
```

```
Use the force [--force|-f] option to suppress this warning.
```

```
Sending format operation ...
```

```
Success formatting namespace:1
```

Then I remade the RAID and FS on top:

```
openqaworker1:~ # mdadm --create /dev/md/openqa --level=0 --raid-devices=2 --run /dev/nvme?n1
```

```
mdadm: Defaulting to version 1.2 metadata
```

```
mdadm: array /dev/md/openqa started
```

```
openqaworker1:~ # mkfs.ext4 /dev/md/openqa
```

```
mke2fs 1.43.8 (1-Jan-2018)
```

```
Geräteblöcke werden verworfen: erledigt
```

```
Ein Dateisystem mit 347662848 (4k) Blöcken und 86917120 Inodes wird erzeugt.
```

```
UUID des Dateisystems: 4dec5b35-616d-420e-93e4-87392ab5d940
```

```
Superblock-Sicherungskopien gespeichert in den Blöcken:
```

```
32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
```

```
4096000, 7962624, 11239424, 20480000, 23887872, 71663616, 78675968,
```

```
102400000, 214990848
```

```
beim Anfordern von Speicher für die Gruppentabellen: erledigt
```

```
Inode-Tabellen werden geschrieben: erledigt
```

```
Das Journal (262144 Blöcke) wird angelegt: erledigt
```

```
Die Superblöcke und die Informationen über die Dateisystemnutzung werden
```

```
geschrieben: erledigt
```

(I had too many tabs open and almost ran that on ow4, whoops. Thankfully mdadm --create couldn't do anything because the devices were busy).

In the t-u shell I edited /etc/fstab to use ext4 for /dev/md/openqa instead of ext2, exited and finally rebooted. It's back up and seems fine so far, but needs to resync the deleted data again. Let's see.

#4 - 2021-12-06 10:44 - maritawerner

- Project changed from openQA Tests to openQA Infrastructure
- Category deleted (Infrastructure)

#5 - 2021-12-06 12:02 - okurz

- Target version set to future

Right now there is a single job running, all others are waiting for assets to download which is still going on.

#6 - 2021-12-07 08:51 - favogt

- Status changed from In Progress to Resolved
- % Done changed from 0 to 100

ow1 failed to come up because of https://bugzilla.opensuse.org/show_bug.cgi?id=1183856 STILL not being fixed on 15.3. I had to make the btrfs snapshot read-write, edit /etc/fstab again, make it read-only and reboot. Should be fine now.

#7 - 2021-12-08 15:55 - dimstar

- Status changed from Resolved to Workable

The worker mis-behaves again, Fabian investigated and it really does look like one of the nVME's needs to be replaced.

#8 - 2021-12-08 16:22 - okurz

- Target version changed from future to Ready

As next steps I suggest to exclude the faulty NVMe from (INTEL SSDPEKNW010T8, Serial Number: BTNH022600SN1P0B) from the /var/lib/openqa partition and we run with remaining NVMe which should still be enough as temporary solution. Then we organize replacement, e.g. create EngInfra ticket, with an invoice forwarded to SUSE LSG QE (if it's not warranty as AFAIR the second NVMe isn't too old).

[favogt](#) feel free to do that or assign to someone from SUSE QE Tools to follow up, e.g. nsinger or me

#9 - 2021-12-08 16:30 - favogt

- Status changed from Workable to Resolved

hdparm -tT /dev/md1 never returned, so the IO issues were back. With dd if=/dev/nvmeXn1 of=/dev/null bs=10M count=10000 I was able to get some performance estimates though: ~1.7GiB/s for nvme0n1, 12-42MiB/s for nvme1n1.

Unfortunately it was bad enough that sync never returned, so a clean shutdown was impossible. In fact, not even reboot -f did anything, the system continued running... I had to force a reset through the BMC (after running sync -f /).

okurz wrote:

As next steps I suggest to exclude the faulty NVMe from (INTEL SSDPEKNW010T8, Serial Number: BTNH022600SN1P0B) from the /var/lib/openqa partition and we run with remaining NVMe which should still be enough as temporary solution. Then we organize replacement, e.g. create EngInfra ticket, with an invoice forwarded to SUSE LSG QE (if it's not warranty as AFAIR the second NVMe isn't too old).

[favogt](#) feel free to do that or assign to someone from SUSE QE Tools to follow up, e.g. nsinger or me

Yep, already did that. mdadm complains when trying to set up a RAID (0 or 1) with just a single disk, so now it's just using /dev/nvme0n1 as ext4 directly.

I had to edit the /etc/fstab of the snapshot itself again for that, meh. It's up and running again, let's see.

#10 - 2021-12-09 08:07 - okurz

- Status changed from Resolved to Feedback

...
okurz wrote:

As next steps I suggest to exclude the faulty NVMe from (INTEL SSDPEKNW010T8, Serial Number: BTNH022600SN1P0B) from the /var/lib/openqa partition and we run with remaining NVMe which should still be enough as temporary solution. Then we organize replacement, e.g. create EngInfra ticket, with an invoice forwarded to SUSE LSG QE (if it's not warranty as AFAIR the second NVMe isn't too old).

[favogt](#) feel free to do that or assign to someone from SUSE QE Tools to follow up, e.g. nsinger or me

Yep, already did that. mdadm complains when trying to set up a RAID (0 or 1) with just a single disk, so now it's just using /dev/nvme0n1 as ext4 directly.

I had to edit the /etc/fstab of the snapshot itself again for that, meh. It's up and running again, let's see.

We commonly create a RAID0 with a single device anyway, see

https://github.com/os-autoinst/salt-states-openqa/blob/master/openqa/nvme_store/openqa-establish-nvme-setup.sh#L33

You set the ticket to "Resolved". I assume that was by mistake because if we have one broken device as long as it's there it's not really fixed, right? So setting to "Feedback" instead for now.

#11 - 2021-12-09 09:19 - maritawerner

[okurz](#) to my knowledge QE LSG has bought most of the openSUSE HW, so I think that the QE Tools team should have a look into that. Or do I miss something here? Check warranty and so on.
[mgriessmeier](#), is that HW that we bought?

#12 - 2021-12-10 10:18 - cdywan

- Assignee changed from favogt to nicksinger

- Priority changed from Urgent to High

Assigning to Nick, to order the new NVME drive after discussing it briefly in the weekly

#13 - 2021-12-10 10:21 - okurz

- Status changed from Feedback to In Progress

maritawerner wrote:

[okurz](#) to my knowledge QE LSG has bought most of the openSUSE HW, so I think that the QE Tools team should have a look into that. Or do I miss something here? Check warranty and so on.

Yes, that is correct. The ticket is already on our backlog (Target version "Ready"), favogt did the initial measurements and workarounds to run of one NVMe and we tracked that. With that workaround in place now nicksinger takes over to order replacement and then delegate to EngInfra to have that replaced within the machine.

#14 - 2021-12-10 10:23 - cdywan

- Related to action #103581: Many jobs on openqa.opensuse.org incomplete in 'timeout: setup exceeded MAX_SETUP_TIME' added

#15 - 2021-12-11 04:10 - openqa_review

- Due date set to 2021-12-25

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

#16 - 2021-12-17 11:21 - mgriessmeier

maritawerner wrote:

[okurz](#) to my knowledge QE LSG has bought most of the openSUSE HW, so I think that the QE Tools team should have a look into that. Or do I miss something here? Check warranty and so on.
[mgriessmeier](#), is that HW that we bought?

not 100% sure, but I assume yes.

Nick is creating the jira-sd ticket right now and I will initiate the ordering.

#17 - 2021-12-17 11:24 - nicksinger

- Status changed from In Progress to Feedback

Requested replacement ordering in <https://sd.suse.com/servicedesk/customer/portal/1/SD-70873>

#18 - 2021-12-22 09:14 - okurz

- Due date changed from 2021-12-25 to 2022-01-28

Was the order actually followed up? Because <https://sd.suse.com/servicedesk/customer/portal/1/SD-70873> looks like Jana Fiserova only closed the ticket after "guidance was provided"

#19 - 2021-12-23 09:33 - nicksinger

I will remind Matthias once again to order the drive

#20 - 2021-12-26 13:24 - okurz

- Description updated

After problems reported by dimstar in <https://suse.slack.com/archives/C02CANHLANP/p1640344725011900> I decreased the worker cache size on openqaworker1 to 100GB and disabled openQA worker instances 16..20. Please follow the "rollback actions" defined in the ticket description after the hardware replacement.

#21 - 2022-01-19 11:08 - okurz

nicksinger informed us from private conversation that we have replacement at the post office in NBG and mgriessmeier needs to pick it up and hand it to SUSE-IT (duh) and then they can continue.

#22 - 2022-01-27 09:01 - mgriessmeier

okurz wrote:

nicksinger informed us from private conversation that we have replacement at the post office in NBG and mgriessmeier needs to pick it up and hand it to SUSE-IT (duh) and then they can continue.

Hi,

sorry the whole delivery mess was my fault - as I was not able to pick it up in time. Meanwhile it was sent back to the vendor and I re-issued delivery to my home address - it is expected to be delivered today. On monday I will be in the office and hand it over to englInfra. Sorry again for delay!

cheers,
Matthias

#23 - 2022-01-31 09:21 - apappas

I am not aware of the current status of the machine and the new SSD so sorry if I am complaining both too late and prematurely but I was asked to do so.

As part of the statistical investigation [requested](#) for a VR, I could force the openqaworker1 too fail or incomplete jobs very consistently.

Here are the results of my VR

https://openqa.opensuse.org/tests/overview?result=failed&result=incomplete&arch=&flavor=&machine=&test=&modules=&module_re=&distri=opensuse&version=15.3&build=vr_13808#

As you can see openqaworker1 takes more than 3 hours to bring a Leap 15.3 system from GM to latest and sometimes just crashes. This does not happen with the other workers.

This has been happening [before Christmas](#) and has been blocking my PR.

#24 - 2022-01-31 10:43 - okurz

- Description updated

Thanks for your update. mkittler has stopped and disabled all workers but 1 to 6 and the containerized s390x workers to reduce the load.

#25 - 2022-01-31 16:16 - nicksinger

- Status changed from Feedback to Blocked

We've placed the new disk for infra and I created <https://sd.suse.com/servicedesk/customer/portal/1/SD-75112> to ask for help building it in.

#26 - 2022-01-31 16:20 - nicksinger

- Due date changed from 2022-01-28 to 2022-02-04

#27 - 2022-02-04 10:43 - cdywan

- Due date changed from 2022-02-04 to 2022-02-11

Bumping the due date due to Will be done on next office day on <https://sd.suse.com/servicedesk/customer/portal/1/SD-75112>

#28 - 2022-02-09 11:20 - favogt

- Status changed from Blocked to Workable
- Assignee changed from nicksinger to favogt
- % Done changed from 100 to 70

Looks like the NVMe got installed \o/

```
openqaworker1:~ # smartctl -a /dev/nvme1n1
smartctl 7.0 2019-05-21 r4917 [x86_64-linux-5.3.18-150300.59.46-default] (SUSE RPM)
Copyright (C) 2002-18, Bruce Allen, Christian Franke, www.smartmontools.org

=== START OF INFORMATION SECTION ===
Model Number:          INTEL SSDPEKNW010T8
Serial Number:         BTNH117605HA1POB
Firmware Version:      004C
PCI Vendor/Subsystem ID: 0x8086
IEEE OUI Identifier:   0x5cd2e4
Controller ID:         1
Number of Namespaces:  1
Namespace 1 Size/Capacity: 1.024.209.543.168 [1,02 TB]
Namespace 1 Formatted LBA Size: 512
Local Time is:         Wed Feb  9 12:18:42 2022 CET
Firmware Updates (0x14): 2 Slots, no Reset required
Optional Admin Commands (0x0017): Security Format Frmw_DL Self_Test
Optional NVM Commands (0x005f):  Comp Wr_Unc DS_Mngmt Wr_Zero Sav/Sel_Feat Timestmp
Maximum Data Transfer Size: 32 Pages
Warning Comp. Temp. Threshold: 77 Celsius
Critical Comp. Temp. Threshold: 80 Celsius
```

Supported Power States

St Op	Max	Active	Idle	RL	RT	WL	WT	Ent_Lat	Ex_Lat
0 +	4.00W	-	-	0	0	0	0	0	0
1 +	3.00W	-	-	1	1	1	1	0	0
2 +	2.20W	-	-	2	2	2	2	0	0
3 -	0.0300W	-	-	3	3	3	3	5000	5000
4 -	0.0040W	-	-	4	4	4	4	5000	9000

Supported LBA Sizes (NSID 0x1)

Id	Fmt	Data	Metadt	Rel_Perf
0 +		512	0	0

=== START OF SMART DATA SECTION ===
SMART overall-health self-assessment test result: PASSED

```
SMART/Health Information (NVMe Log 0x02)
Critical Warning:          0x00
Temperature:              33 Celsius
Available Spare:          100%
Available Spare Threshold: 10%
Percentage Used:          0%
Data Units Read:          9 [4,60 MB]
Data Units Written:       0
Host Read Commands:      216
Host Write Commands:     0
Controller Busy Time:    0
Power Cycles:             1
Power On Hours:          28
Unsafe Shutdowns:        0
Media and Data Integrity Errors: 0
Error Information Log Entries: 0
Warning Comp. Temperature Time: 0
Critical Comp. Temperature Time: 0
```

Error Information (NVMe Log 0x01, max 256 entries)
No Errors Logged

The system is currently idle, so I stopped the workers on it and will recreate the RAID after lunch and increase the worker count afterwards again.

#29 - 2022-02-09 15:43 - favogt

- Status changed from Workable to Resolved

- % Done changed from 70 to 100

```
openqaworker1:~ # mdadm --create /dev/md/openqa --level=0 --raid-devices=2 --run /dev/nvme?n1
mdadm: /dev/nvme0n1 appears to contain an ext2fs file system
      size=390711384K  mtime=Wed Feb  9 03:33:46 2022
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md/openqa started.
openqaworker1:~ # mdadm --stop /dev/md/openqa
mdadm: stopped /dev/md/openqa
openqaworker1:~ # wipefs -a /dev/nvme0n1
/dev/nvme0n1: 4 Bytes wurden an Position 0x00001000 gelöscht (linux_raid_member): fc 4e 2b a9
/dev/nvme0n1: 2 Bytes wurden an Position 0x00000438 gelöscht (ext4): 53 ef
openqaworker1:~ # mdadm --create /dev/md/openqa --level=0 --raid-devices=2 --run /dev/nvme?n1
mdadm: /dev/nvme1n1 appears to be part of a raid array:
      level=raid0 devices=2 ctime=Wed Feb  9 16:21:24 2022
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md/openqa started.
```

Then I set up the mount in a transactional-update shell:

```
openqaworker1:~ # transactional-update shell
Checking for newer version.
transactional-update 3.2.2 started
Options: shell
Separate /var detected.
2022-02-09 16:23:08 tukit 3.2.2 started
2022-02-09 16:23:08 Options: -c1221 open
2022-02-09 16:23:09 Using snapshot 1221 as base for new snapshot 1222.
2022-02-09 16:23:09 Syncing /etc of previous snapshot 1220 as base into new snapshot /.snapshots/1222/snapshot
ID: 1222
2022-02-09 16:23:12 Transaction completed.
Opening chroot in snapshot 1222, continue with 'exit'
2022-02-09 16:23:12 tukit 3.2.2 started
2022-02-09 16:23:12 Options: call 1222 bash
2022-02-09 16:23:12 Executing `bash`:
transactional update # mkfs.ext4 /dev/md/openqa
mke2fs 1.43.8 (1-Jan-2018)
Discarding device blocks: done
Creating filesystem with 347662848 4k blocks and 86917120 inodes
Filesystem UUID: f8df103a-c2da-4f5e-b1aa-2e9064bcbe6a
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424, 20480000, 23887872, 71663616, 78675968,
    102400000, 214990848

Allocating group tables: done
Writing inode tables: done
Creating journal (262144 blocks): done
Writing superblocks and filesystem accounting information: done

transactional update # nano /etc/fstab
transactional update # mount /var/lib/openqa
transactional update # umount /var/lib/openqa
transactional update # exit
2022-02-09 16:26:03 Application returned with exit status 0.
```

To work around the /etc/fstab handling bug in transactional-update, I copied the fstab from the resulting snapshot overlay into the snapshot itself:

```
openqaworker1:~ # btrfs prop set /.snapshots/1222/snapshot ro false
openqaworker1:~ # diff -u /var/lib/overlay/1222/etc/fstab /.snapshots/1222/snapshot/etc/fstab
--- /var/lib/overlay/1222/etc/fstab      2022-02-09 16:24:39.869538443 +0100
+++ /.snapshots/1222/snapshot/etc/fstab  2022-02-09 16:23:10.448519218 +0100
@@ -1,6 +1,6 @@
 UUID=ff1922d2-d2e4-4860-9634-acac681dd0f9 / btrfs ro 0 0
 UUID=ff1922d2-d2e4-4860-9634-acac681dd0f9 /var btrfs subvol=/@/var,x-initrd.mount 0 0
-/dev/md/openqa /var/lib/openqa ext4 defaults 0 0
+/dev/nvme0n1 /var/lib/openqa ext4 defaults 0 0
 UUID=ff1922d2-d2e4-4860-9634-acac681dd0f9 /usr/local btrfs subvol=/@/usr/local 0 0
 UUID=ff1922d2-d2e4-4860-9634-acac681dd0f9 /tmp btrfs subvol=/@/tmp 0 0
 UUID=ff1922d2-d2e4-4860-9634-acac681dd0f9 /srv btrfs subvol=/@/srv 0 0
openqaworker1:~ # cp /var/lib/overlay/1222/etc/fstab /.snapshots/1222/snapshot/etc/fstab
openqaworker1:~ # btrfs prop set /.snapshots/1222/snapshot ro true
```

After a reboot it came up properly, and I unmasked all openqa-worker-auto-restart@ services and enabled workers 1-12 again.

Let's see how it goes!

#30 - 2022-02-09 16:24 - okurz

Thank you. I double checked and all looks good from my side as well.

#31 - 2022-02-10 10:58 - favogt

How many and which workers did ow1 have configured originally? IIRC it went up to 20, but there might've been a gap. 1-12 are configured to not have qemu_i586,vagrant currently.

#32 - 2022-02-10 21:37 - okurz

favogt wrote:

How many and which workers did ow1 have configured originally? IIRC it went up to 20, but there might've been a gap. 1-12 are configured to not have qemu_i586,vagrant currently.

As the description says before this ticket we had 16 instances.

#33 - 2022-02-14 13:03 - favogt

okurz wrote:

favogt wrote:

How many and which workers did ow1 have configured originally? IIRC it went up to 20, but there might've been a gap. 1-12 are configured to not have qemu_i586,vagrant currently.

As the description says

Indeed, totally missed that...

before this ticket we had 16 instances.

Actually, it mentions [7..16] as well as {16..20}. I enabled 1-20 now, let's see. Will be a good test for the new NVMe at least!

#34 - 2022-02-17 14:01 - favogt

- Copied to action #107017: Random asset download (cache service) failures on openqaworker1 added

#35 - 2022-02-18 10:20 - favogt

- Has duplicate action #107017: Random asset download (cache service) failures on openqaworker1 added

#36 - 2022-02-18 10:26 - favogt

As suggested by bmwiedemann, I ran fstrim -v /var/lib/openqa to discard all unused blocks. This took a while but brought (read!) performance back again.

I changed fstrim.timer to run daily by adding /etc/systemd/system/fstrim.timer.d/daily.conf:

```
[Unit]
Description=Discard unused blocks once a day

[Timer]
OnCalendar=daily
```

Currently the system is running 20 jobs and both hdparm and dd report acceptable speed:

```
openqaworker1:~ # ps -ef | grep qemu | wc -l
21
openqaworker1:~ # hdparm --direct -t -T /dev/md127

/dev/md127:
Timing O_DIRECT cached reads: 1484 MB in 2.00 seconds = 741.80 MB/sec
Timing O_DIRECT disk reads: 6608 MB in 3.00 seconds = 2202.56 MB/sec
openqaworker1:~ # dd if=/dev/nvme1n1 bs=1M count=1000 status=progress iflag=direct of=/dev/null
1000+0 Datensätze ein
```

```
1000+0 Datensätze aus
1048576000 Bytes (1,0 GB, 1000 MiB) kopiert, 0,763541 s, 1,4 GB/s
openqaworker1:~ # dd if=/dev/nvme0n1 bs=1M count=1000 status=progress iflag=direct of=/dev/null
1000+0 Datensätze ein
1000+0 Datensätze aus
1048576000 Bytes (1,0 GB, 1000 MiB) kopiert, 0,702554 s, 1,5 GB/s
```

#37 - 2022-02-18 10:28 - kraih

- Status changed from Resolved to New

This issue is not resolved, as [#107017](#) has shown.

#38 - 2022-02-18 10:32 - kraih

- Priority changed from High to Urgent

#39 - 2022-02-18 10:39 - favogt

kraih wrote:

This issue is not resolved, as [#107017](#) has shown.

After running fstrim it's currently "resolved again". Did an error appear after the current boot?

#40 - 2022-02-21 08:26 - okurz

- Due date changed from 2022-02-11 to 2022-02-24

The performance looks ok

```
# hdparm --direct -t -T /dev/md127
```

```
/dev/md127:
```

```
Timing O_DIRECT cached reads: 1560 MB in 2.00 seconds = 779.75 MB/sec
Timing O_DIRECT disk reads: 4470 MB in 3.04 seconds = 1469.18 MB/sec
```

No related errors in dmesg. 20 jobs running. However I think 20 worker instances is too much. For example there are incompletes like <https://openga.opensuse.org/tests/2200039> with "cache failure: Cache service queue already full (5)" which likely mean that we can't run that many right now unless we also tweak parameters about the worker cache.

#41 - 2022-02-21 10:11 - kraih

okurz wrote:

The performance looks ok

```
# hdparm --direct -t -T /dev/md127
```

```
/dev/md127:
```

```
Timing O_DIRECT cached reads: 1560 MB in 2.00 seconds = 779.75 MB/sec
Timing O_DIRECT disk reads: 4470 MB in 3.04 seconds = 1469.18 MB/sec
```

Actually, the machine might be in trouble again. Running `hdparm --direct -t -T /dev/md127` hangs now.

#42 - 2022-02-21 10:15 - favogt

kraih wrote:

okurz wrote:

The performance looks ok

```
# hdparm --direct -t -T /dev/md127
```

```
/dev/md127:
```

```
Timing O_DIRECT cached reads: 1560 MB in 2.00 seconds = 779.75 MB/sec
Timing O_DIRECT disk reads: 4470 MB in 3.04 seconds = 1469.18 MB/sec
```

Actually, the machine might be in trouble again. Running `hdparm --direct -t -T /dev/md127` hangs now.

Yup, dd is also slow again. fstrim ran 10h ago, which should be recent enough, though it's theoretically possible that it's been written full during that time period already. If another fstrim helps, the discard FS option is probably worth a try.

#43 - 2022-02-21 10:16 - okurz

- Subject changed from *OW1: performance loss* to *OW1: performance loss size:M*
- Description updated
- Status changed from *New* to *Workable*

#44 - 2022-02-22 10:05 - mkittler

@fvogt says after another fstrim the performance looks still ok and he added the discard FS option

#45 - 2022-02-24 13:03 - mkittler

- Status changed from *Workable* to *Feedback*

So, are we good with that measures in place?

#46 - 2022-02-25 11:09 - favogt

Performance is still looking ok, but there's some random corruption in some downloaded assets, causing e.g. downloaded HDDs to fail booting or upgrading:

https://openqa.opensuse.org/tests/2209518#step/upgrade_select/4

#47 - 2022-03-04 10:32 - okurz

- Due date deleted (2022-02-24)
- Target version changed from *Ready to future*

It looks serious enough so we would like us to handle this seriously.

[favogt](#) I will leave this to you decide if there are further actions needed. I am excluding the ticket from our backlog and remove the due date because we will only react on this on your request. If you can provide a good proof that the hardware is at fault we can order replacement (again).

#48 - 2022-03-04 10:42 - favogt

- Assignee deleted (*favogt*)

Except for the occasional bad download every few days, the system appears to work fine to me.

I don't know what it's caused by and I'm not sure what the best way to find out is.

#49 - 2022-03-09 09:07 - okurz

- Status changed from *Feedback* to *Workable*
- Priority changed from *Urgent* to *High*
- Target version changed from *future* to *Ready*

so back to the team for checking again

#50 - 2022-03-09 10:44 - cdywan

We do have two NVME devices, and we should experiment here, such as by trying out the new NVME and collecting figures from before and after (avoid looking at too old jobs)

#51 - 2022-03-11 11:13 - okurz

I suggest to use <https://progress.opensuse.org/projects/openqatests/wiki/Wiki#Statistical-investigation> , run some tests without changes, then change a parameter and test again

#52 - 2022-03-24 15:40 - mkittler

Unlike AC1 might let one assume, ow1 is still used in production. In fact, it seems to be the worker processing the most jobs.

It doesn't look like its failure/incomplete rate is significantly worse compared to the other workers:

```
openqa=> with finished as (select result, t_finished, host from jobs left join workers on jobs.assigned_worker_id = workers.id where result != 'none') select host, round(count(*) filter (where result='failed' or result='
```

```
incomplete') * 100. / count(*), 2)::numeric(5,2)::float as ratio_failed_by_host, count(*) total from finished
where t_finished >= '2022-03-04' group by host order by ratio_failed_by_host desc;
```

host	ratio_failed_by_host	total
siodtw01	74.89	223
openqaworker1_container	74.88	211
openqa-aarch64	53.37	5134
ip-10-252-32-90	37.92	654
ip-10-252-32-98	36.52	2141
power8	33.92	2621
imagetester	27.87	854
openqaworker1	27.22	8599
openqaworker4	26.64	7669
openqaworker7	23.89	6860
	1.44	2785

(11 Zeilen)

I'm ignoring openqaworker1_container here because that's likely not a representative set of jobs.

#53 - 2022-03-24 21:23 - okurz

- Status changed from *Workable* to *Resolved*

- Assignee set to *mkittler*

mkittler wrote:

Unlike AC1 might let one assume, ow1 is still used in production. In fact, it seems to be the worker processing the most jobs.

That is correct. I wondered if maybe not all worker instances are enabled but the machine seems to be happily running 20 openQA worker instances so back to the all time high :)

It doesn't look like its failure/incomplete rate is significantly worse compared to the other workers:

[...]

That looks good. Thanks for crosschecking. I also crosschecked the raid config and everything looks to be ok. I would say that's enough then. So resolving the ticket and assigning mkittler as the main person driving this story to completion.