
Subject: Re: OpenVZ 7 containers crashing with ext4 errors

Posted by [allan.talver](#) on Fri, 11 Sep 2020 14:08:50 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello,

Wanted to give a short update and ask a couple of questions.

First of all, last night one of the virtual containers on a non-patched host went into read-only. However, different from the previous cases is that pcompact was not involved in this case. We actually have pcompact cron disabled and we trigger it manually on the nodes that are in our test sample. The errors in messages log were:

Sep 11 00:34:00 server-n697 kernel: bash (639778): drop_caches: 3

Sep 11 00:34:36 server-n697 systemd: Started Session c185489 of user root.

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_free_blocks:4933: Out of memory

Sep 11 00:34:39 server-n697 kernel: Aborting journal on device ploop35478p1-8.

Sep 11 00:34:39 server-n697 kernel: EXT4-fs (ploop35478p1): Remounting filesystem read-only

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_ext_remove_space:3073: IO failure

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_ext_truncate:4692: Journal has aborted

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_reserve_inode_write:5358: Journal has aborted

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_truncate:4145: Journal has aborted

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_reserve_inode_write:5358: Journal has aborted

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_orphan_del:2731: Journal has aborted

Sep 11 00:34:39 server-n697 kernel: EXT4-fs error (device ploop35478p1) in ext4_reserve_inode_write:5358: Journal has aborted

Sep 11 00:34:39 server-n697 kernel: EXT4-fs (ploop35478p1): ext4_writepages: jbd2_start: 0 pages, ino 661650; err -30

Sep 11 00:34:41 server-n697 kernel: dd invoked oom-killer: gfp_mask=0xd0, order=0, oom_score_adj=0

Sep 11 00:34:41 server-n697 kernel: dd cpuset=3988 mems_allowed=0

Sep 11 00:34:41 server-n697 kernel: CPU: 6 PID: 273822 Comm: dd ve: 3988 Kdump: loaded Not tainted 3.10.0-1127.8.2.vz7.151.14 #1 151.14

Sep 11 00:34:41 server-n697 kernel: Hardware name: Supermicro X9DRW/X9DRW, BIOS 3.0c 03/24/2014

Sep 11 00:34:41 server-n697 kernel: Call Trace:

Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb95b67f1>] dump_stack+0x19/0x1b

Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb95b0fc6>] dump_header+0x90/0x229

Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb8fd7076>] ? find_lock_task_mm+0x56/0xc0

Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb8fd7dad>] oom_kill_process+0x47d/0x640

Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb90040fe>] ? get_task_oom_score_adj+0xee/0x100

Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb8fd7213>] ? oom_badness+0x133/0x1e0

```

Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb905f509>]
mem_cgroup_oom_synchronize+0x4b9/0x510
Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb8fd84c3>] pagefault_out_of_memory+0x13/0x50
Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb95af06d>] mm_fault_error+0x6a/0x157
Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb95c49a1>] __do_page_fault+0x491/0x500
Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb95c4a45>] do_page_fault+0x35/0x90
Sep 11 00:34:41 server-n697 kernel: [<ffffffffffb95c0778>] page_fault+0x28/0x30
Sep 11 00:34:41 server-n697 kernel: Task in /machine.slice/3988 killed as a result of limit of
/machine.slice/3988
Sep 11 00:34:41 server-n697 kernel: memory: usage 4192232kB, limit 4194304kB, failcnt
13429301
Sep 11 00:34:41 server-n697 kernel: memory+swap: usage 4325376kB, limit 4325376kB, failcnt
31733768989
Sep 11 00:34:41 server-n697 kernel: kmem: usage 36356kB, limit 9007199254740988kB, failcnt 0
Sep 11 00:34:41 server-n697 kernel: Memory cgroup stats for /machine.slice/3988:
rss_huge:479232KB mapped_file:33780KB shmem:84KB slab_unreclaimable:9776KB
swap:133144KB cache:3459360KB rss:696520KB slab_reclaimable:10664KB
inactive_anon:199280KB active_anon:497324KB inactive_file:597760KB active_file:2861360KB
unevictable:0KB
Sep 11 00:34:41 server-n697 kernel: Memory cgroup out of memory: Kill process 671560
(mysql) score 87 or sacrifice child
Sep 11 00:34:41 server-n697 kernel: Killed process 581188 (mysql) in VE "3988", UID 116,
total-vm:2820740kB, anon-rss:376664kB, file-rss:0kB, shmem-rss:0kB
Sep 11 00:34:53 server-n697 systemd: Started Session c185490 of user root.
Sep 11 00:35:01 server-n697 kernel: bash (639778): drop_caches: 3

```

Please note that the container was running the script that was creating and deleting random files. The script we use to generate read-write activity for testing purposes. Of course, the script also failed after filesystem was switched to read-only mode.

Would you think that the cause of this error is related to the same bug that the kernel patch should fix? Meaning that pcompact is not part of the problem, but just happens to amplify the issue.

Secondly, we see that while pcompact is running, the virtual container disk utilisation fluctuates rapidly and the disk also comes 100% full several time. Is that an expected behaviour during a pcompact run? As a result, we have seen some applications (like MySQL) throwing errors saying that it's unable to write to disk. Feels like it could cause data corruption.

Thirdly, worth noting that we have not yet seen a container switching to read-only on a node that has the kernel patch applied. But even on that node, the disk was showing as full when pcompact ran, so the concern regarding potential data corruption applies.