Subject: OpenVZ 7 containers crashing with ext4 errors
Posted by allan.talver on Fri, 03 Jul 2020 15:06:35 GMT
View Forum Message <> Reply to Message

We have been running OpenVZ 6 in our environments for several years. The platform has been stable and predictable. Recently we have started evaluating OpenVZ 7 as the replacement. In most aspects, OpenVZ 7 has proven to be good and suitable for our purpose. However, recently we began to experience seemingly random crashes with symptoms pointing to ext4 filesystem and ploop. When the crash happens, virtual container is left with its disk in read only state. Restart is not successful due to errors present in the filesystem. After running fsck manually, container is able to start and we have not experienced any data loss. However, even without data loss, such events reduce the confidence to run production workloads with critical data on these servers.

In total we have now had 4 such events. First 3 were Ubuntu 16.04 containers that got migrated from OpenVZ 6 to OpenVZ 7. Before migration with ovztransfer.sh, the server disks got converted from simfs to ploop with vzctl convert. Initially we thought that the issue might be something in our migration procedure or something specific to Ubuntu 16.04 operating system, because no Ubuntu 18.04 server (created fresh on OpenVZ 7) had crashed. However, two days ago the container affected by the last crash was 18.04 which never was migrated from OpenVZ 6 (although it has been migrated between OpenVz 7 nodes).

Another thing we have noticed is that the crash seems to be happening roughly at the same time when pcompact is running on the hardware node. And also, 2 out of 3 containers that had been migrated from OpenVZ 6, crashed the next night right after the migration.

Are these errors something that the community has seen before and could help us explain?

In all cases, the log output in hardware node dmesg has been similar and as follows:

[2020-05-29 02:02:20] WARNING: CPU: 12 PID: 317821 at fs/ext4/ext4_jbd2.c:266 __ext4_handle_dirty_metadata+0x1c2/0x220 [ext4]
[2020-05-29 02:02:20] Modules linked in: nfsv3 nfs_acl ip6table_mangle nf_log_ipv4 nf_log_common xt_LOG nfsv4 dns_resolver nfs lockd grace fscache xt_multiport nf_conntrack_ipv4 nf_defrag_ipv4 xt_conntrack xt_comment binfmt_misc xt_CHECKSUM iptable_mangle ip6t_REJECT nf_reject_ipv6 ipt_REJECT nf_reject_ipv4 tun 8021q garp mrp devlink ip6table_filter ip6_tables iptable_filter bonding ebtable_filter ebt_among ebtables sunrpc iTCO_wdt iTCO_vendor_support sb_edac intel_powerclamp coretemp intel_rapl iosf_mbi kvm_intel kvm irqbypass crc32_pclmul ghash_clmulni_intel aesni_intel lrw gf128mul glue_helper ablk_helper cryptd pcspkr i2c_i801 joydev mei_me lpc_ich mei sg ioatdma wmi ipmi_si ipmi_devintf ipmi_msghandler acpi_power_meter acpi_pad pcc_cpufreq ip_vs nf_conntrack libcrc32c br_netfilter veth overlay ip6_vzprivnet
[2020-05-29 02:02:20] ip6_vznetstat ip_vznetstat ip_vzprivnet vziolimit vzevent vzlist vzstat vznetstat vznetdev vzmon vzdev bridge stp llc ip_tables ext4 mbcache jbd2 sd_mod crc_t10dif crct10dif_generic ast i2c_algo_bit drm_kms_helper syscopyarea sysfillrect sysimgblt fb_sys_fops ttm ixgbe ahci drm libahci libata megaraid_sas crct10dif_pclmul crct10dif_common crc32c_intel mdio ptp pps_core drm_panel_orientation_quirks dca dm_mirror dm_region_hash dm_log dm_mod pio_kaio pio_nfs pio_direct pfmt_raw pfmt_ploop1 ploop
[2020-05-29 02:02:20] CPU: 12 PID: 317821 Comm: e4defrag2 ve: 0 Kdump: loaded Not tainted

3.10.0-1062.12.1.vz7.131.10 #1 131.10
[2020-05-29 02:02:20] Hardware name: Supermicro SYS-1028R-WC1RT/X10DRW-iT, BIOS 2.0a 07/26/2016
[2020-05-29 02:02:20] Call Trace:
[2020-05-29 02:02:20] [<ffffffff81baebc7>] dump_stack+0x19/0x1b
[2020-05-29 02:02:20] [<ffffffff8149bdc8>] __warn+0xd8/0x100
[2020-05-29 02:02:20] [<ffffffff8149bf0d>] warn_slowpath_null+0x1d/0x20
[2020-05-29 02:02:20] [<ffffffffc047cd82>] __ext4_handle_dirty_metadata+0x1c2/0x220 [ext4]
[2020-05-29 02:02:20] [<ffffffffc0474f44>] ext4_ext_split+0x304/0x9a0 [ext4]
[2020-05-29 02:02:20] [<ffffffffc0476dfd>] ext4_ext_insert_extent+0x7bd/0x8d0 [ext4]
[2020-05-29 02:02:20] [<ffffffffc0479a7f>] ext4_ext_map_blocks+0x5cf/0xf60 [ext4]
[2020-05-29 02:02:20] [<ffffffffc0446676>] ext4_map_blocks+0x136/0x6b0 [ext4]
[2020-05-29 02:02:20] [<ffffffffc047496c>] ? ext4_alloc_file_blocks.isra.36+0xbc/0x2f0 [ext4]
[2020-05-29 02:02:20] [<ffffffffc047498f>] ext4_alloc_file_blocks.isra.36+0xdf/0x2f0 [ext4]
[2020-05-29 02:02:20] [<ffffffffc047b6bd>] ext4_fallocate+0x15d/0x990 [ext4]
[2020-05-29 02:02:20] [<ffffffff8166cb78>] ? __sb_start_write+0x58/0x120
[2020-05-29 02:02:20] [<ffffffff81666c72>] vfs_fallocate+0x142/0x1e0
[2020-05-29 02:02:20] [<ffffffff81667cdb>] SyS_fallocate+0x5b/0xa0
[2020-05-29 02:02:20] [<ffffffff81bc1fde>] system_call_fastpath+0x25/0x2a
[2020-05-29 02:02:20] ---[end trace cf8fe0ecbf57efcc]---
[2020-05-29 02:02:20] EXT4-fs: ext4_ext_split:1139: aborting transaction: error 28 in __ext4_handle_dirty_metadata
[2020-05-29 02:02:20] EXT4-fs error (device ploop52327p1): ext4_ext_split:1139: inode #325519: block 6002577: comm e4defrag2: journal_dirty_metadata failed: handle type 3 started at line 4741, credits 8/0, errcode -28
[2020-05-29 02:02:20] Aborting journal on device ploop52327p1-8.
[2020-05-29 02:02:20] EXT4-fs (ploop52327p1): Remounting filesystem read-only
[2020-05-29 02:02:20] EXT4-fs error (device ploop52327p1) in ext4_free_blocks:4915: Journal has aborted
[2020-05-29 02:02:20] EXT4-fs error (device ploop52327p1) in ext4_free_blocks:4915: Journal has aborted
[2020-05-29 02:02:20] EXT4-fs error (device ploop52327p1) in ext4_reserve_inode_write:5360: Journal has aborted
[2020-05-29 02:02:20] EXT4-fs error (device ploop52327p1) in ext4_alloc_file_blocks:4753: error 28

Thanks!