Subject: Re: Processes in D state when vzctl chkpnt suspend
Posted by Andrew Vagin on Wed, 21 Mar 2012 12:21:12 GMT
View Forum Message <> Reply to Message

On 03/20/2012 08:44 PM, Stoyan Stoyanov wrote:
> Hi,
>
> I have an issue when trying vzbackups that happens randomly.
> The issue is with the vzctl chkpnt veid --suspend .
>
> what happens is , all ve's process goes into D states.
> no logs on dmesg or anywhere on the node system in the container itself.
> As you know these processes are uninterruptible (un-killable).
> I'm not sure what exactly happens, so please help me.
> vzserver doesn't use nfs or something like that, but fs is on lvms.
> the kernel version is: Linux vz2 2.6.32-5-openvz-amd64 #1 SMP Mon Oct
> 3 05:12:50 UTC 2011 x86_64 GNU/Linux
I recommend you to use our rhel6-2.6.32 kernel.
http://download.openvz.org/kernel/branches/rhel6-2.6.32/
>
> here are the ps axu output from the node, only for the freezed
> container processes.:
> 204 root 6688 0.0 0.0 8352 636 ? Ds Mar12 0:01
> init [2]
> 204 root 7296 0.0 0.0 119692 1292 ? Dl Mar12 0:01
> /usr/sbin/rsyslogd -c4
> 204 root 7366 0.0 0.0 82588 3316 ? Ds Mar12 0:12
> /usr/sbin/apache2 -k start
> 204 root 7384 0.0 0.0 20900 712 ? Ds Mar12 0:01
> /usr/sbin/cron
> 204 root 7577 0.0 0.0 37160 2096 ? Ds Mar12 0:00
> /usr/lib/postfix/master
> 204 101 7587 0.0 0.0 39380 2224 ? D Mar12 0:00
> qmgr -l -t fifo -u
> 204 root 7622 0.0 0.0 49168 960 ? Ds Mar12 0:00
> /usr/sbin/sshd
> 204 101 8899 0.0 0.0 39224 2132 ? D Mar17 0:00
> pickup -l -t fifo -u -c
> 204 www-data 25719 0.0 0.0 82728 4044 ? D Mar17 0:00
> /usr/sbin/apache2 -k start
> 204 www-data 26052 0.0 0.0 82728 4032 ? D Mar17 0:00
> /usr/sbin/apache2 -k start
> 204 www-data 26894 0.0 0.0 82728 3900 ? D Mar17 0:00
> /usr/sbin/apache2 -k start
> 204 www-data 27409 0.0 0.0 82728 3860 ? D Mar17 0:00
> /usr/sbin/apache2 -k start
> 204 www-data 27542 0.0 0.0 82728 3832 ? D Mar17 0:00
> /usr/sbin/apache2 -k start

> 204 www-data 27905 99.6 0.0 82728 3824 ?   R   Mar17 5182:40
> /usr/sbin/apache2 -k start

This process is in RUNNING state... Could you say what it's doing.

strace -fp 3824 -o log.s
cat /proc/3824/stack

> 204 www-data 28113 0.0 0.0 82728 3768 ?   D   Mar17   0:00
> /usr/sbin/apache2 -k start
> 204 www-data 28191 0.0 0.0 82728 3760 ?   D   Mar17   0:00
> /usr/sbin/apache2 -k start
> 204 www-data 28347 0.0 0.0 82728 3708 ?   D   Mar17   0:00
> /usr/sbin/apache2 -k start
> 204 www-data 28720 0.0 0.0 82728 3628 ?   D   Mar17   0:00
> /usr/sbin/apache2 -k start
> 204 www-data 28750 0.0 0.0 82728 3596 ?   D   Mar17   0:00
> /usr/sbin/apache2 -k start
> 204 www-data 28849 0.0 0.0 82728 3560 ?   D   Mar17   0:00
> /usr/sbin/apache2 -k start
> 204 root   28956 99.3 0.0 10220   520 ?   Rs   Mar17 5163:04
> /usr/sbin/vzctl chkpnt 204 --suspend
>
> as you see all of them are in D state.

Not all and it's a problem.

>
> here is the stack trace for the vzctl chkpnt process
>
> [714486.771855] Pid: 28956, comm: vzctl Not tainted
> 2.6.32-5-openvz-amd64 #1 feoktistov X9SCL/X9SCM
> [714486.771857] RIP: 0010:[<ffffffff810484cf>] [<ffffffff810484cf>]
> wait_task_inactive+0x41/0xfb
> [714486.771861] RSP: 0018:ffff8803578f1cf8 EFLAGS: 00000246
> [714486.771863] RAX: 0000000000000001 RBX: 800000000000015d RCX:
> ffff8803578f1c78
> [714486.771864] RDX: ffff880011a56940 RSI: 0000000000000296 RDI:
> 0000000000000292
> [714486.771866] RBP: ffff880421c2e800 R08: ffff8803578f0000 R09:
> ffff88043a160780
> [714486.771868] R10: 0000000100000000 R11: ffff880011b96940 R12:
> ffff880011a56940
> [714486.771869] R13: 0000000000000000 R14: 0000000000016940 R15:
> ffff88043d280800
> [714486.771871] FS: 00007f11a6e7e700(0000) GS:ffff880011b80000(0000)
> knlGS:0000000000000000
> [714486.771873] CS: 0010 DS: 0000 ES: 0000 CR0: 000000008005003b

> [714486.771875] CR2: 00007f9c12391ae0 CR3: 000000041f983000 CR4:
> 00000000000406e0
> [714486.771877] DR0: 0000000000000000 DR1: 0000000000000000 DR2:
> 0000000000000000
> [714486.771878] DR3: 0000000000000000 DR6: 00000000ffff0ff0 DR7:
> 0000000000000400
> [714486.771880] Call Trace:
> [714486.771881] <NMI> <<EOE>> [<ffffffffa03defb6>] ?
> cpt_vps_suspend+0xede/0x138a [vzcpt]
> [714486.771887] [<ffffffffa03dca7f>] ? cpt_ioctl+0x5e5/0xcd2 [vzcpt]
> [714486.771889] [<ffffffffa03dc49a>] ? cpt_ioctl+0x0/0xcd2 [vzcpt]
> [714486.771891] [<ffffffff81134cde>] ? proc_reg_unlocked_ioctl+0xa2/0xc2
> [714486.771894] [<ffffffff810fd096>] ? vfs_ioctl+0x21/0x6c
> [714486.771896] [<ffffffff810fd5d3>] ? do_vfs_ioctl+0x47c/0x4cb
> [714486.771899] [<ffffffff810f1aa4>] ? vfs_write+0xcd/0x102
> [714486.771901] [<ffffffff810fd65f>] ? sys_ioctl+0x3d/0x5c
> [714486.771903] [<ffffffff81010c12>] ? system_call_fastpath+0x16/0x1b
> [714486.771904] Pid: 28956, comm: vzctl Not tainted
> 2.6.32-5-openvz-amd64 #1
> [714486.771905] Call Trace:
> [714486.771906] <NMI> [<ffffffff8100fdda>] ? show_regs+0x3c/0x5d
> [714486.771909] [<ffffffff812ec738>] ? nmi_watchdog_tick+0xb7/0x1aa
> [714486.771912] [<ffffffff812ebe83>] ? do_nmi+0xa5/0x264
> [714486.771914] [<ffffffff812eb920>] ? nmi+0x20/0x30
> [714486.771916] [<ffffffff810484cf>] ? wait_task_inactive+0x41/0xfb
> [714486.771917] <<EOE>> [<ffffffffa03defb6>] ?
> cpt_vps_suspend+0xede/0x138a [vzcpt]
> [714486.771921] [<ffffffffa03dca7f>] ? cpt_ioctl+0x5e5/0xcd2 [vzcpt]
> [714486.771924] [<ffffffffa03dc49a>] ? cpt_ioctl+0x0/0xcd2 [vzcpt]
> [714486.771926] [<ffffffff81134cde>] ? proc_reg_unlocked_ioctl+0xa2/0xc2
> [714486.771928] [<ffffffff810fd096>] ? vfs_ioctl+0x21/0x6c
> [714486.771931] [<ffffffff810fd5d3>] ? do_vfs_ioctl+0x47c/0x4cb
> [714486.771933] [<ffffffff810f1aa4>] ? vfs_write+0xcd/0x102
> [714486.771935] [<ffffffff810fd65f>] ? sys_ioctl+0x3d/0x5c
> [714486.771937] [<ffffffff81010c12>] ? system_call_fastpath+0x16/0x1b
>
> I guess I know what's happen, but I don't know how to fix and I want
> to hear some suggestions.
>
> Is there anyone else that suffer of such issue ?
> Do you have any idea what happens and if I can provide some other
> useful info , please write.
>
>
>
>
>
>

> Stoyan Stoyanov
> Core System Administrator
>
>
>
> CONFIDENTIAL
> The information contained in this email and any attachment is
> confidential. It is intended only for the named addressee(s). If you
> are not the named addressee(s) please notify the sender immediately
> and do not disclose, copy or distribute the contents to any other
> person other than the intended addressee(s).
>
>
>