Subject: Re: Is there a stable OpenVZ kernel, and which should be fit for production
Posted by kir on Tue, 22 Nov 2011 10:34:59 GMT
View Forum Message <> Reply to Message

On 11/22/2011 12:52 PM, Dariush Pietrzak wrote:
> Hello,
> since 2.6.32 branch is no longer maintained:
> "
> Also, from now (30 August 2011) we no longer maintain the following kernel branches:
>
> * 2.6.27
> * 2.6.32
> "
> we have switched to RHEL6 branch, which seems to run fine, and solves some
> long-running problems with 2.6.32 (vSwap, problem with accounting of
> mmaped files usage).
> All was nice until some heavier loaded servers came online with RHEL6, and
> - they started crashing. And then came the upgrade train:
> stab036.1 => stab037.1 => stab039.10 => stab040.1 => stab042.1 etc
>
> With one of the problems we caught, we were told to switch from stable to
> testing kernels (now I see that that testing kernel later became stable,
> so while confusing, it makes some sense).
> All those kernels (and stab039.11, which from description should be
> latest stable) exhibit the same problem/class of problems - when put under
> stress, they crash.
> It's quite easy to recreate, now that we've spent some time tracking it down,
> just start the machine with for example:
>
> stress --cpu 12 --io 16 --vm 32 -d 24 --hdd-bytes 10G
> and maybe bonnie++ running in loop, and in few minutes/few hours you've got
> dead machines spewing something like:
>
> [1515.249585] BUG: scheduling while atomic: stress/2054/0xffff8800
> [1515.250189] BUG: unable to handle kernel paging request at ffffffc047118e0
> [1515.250189] IP: [<ffffffff8105620e>] account_system_time+0x9e/0x1f0
> [1515.250189] PGD 1a27067 PUD 0
> [1515.250189] Thread overran stack, or stack corrupted
> [1515.250189] Oops: 0000 [#1] SMP
>
> or maybe:
> [1876.747809] BUG: unable to handle kernel paging request at 00000006000000bd
> [1876.747815] IP: [<ffffffff8105a4fe>] select_task_rq_fair+0x32e/0xa20
> [1876.747823] PGD 12d089067 PUD 0
> [1876.747826] Oops: 0000 [#1] SMP
>
> or

> [38764.623677] BUG: unable to handle kernel paging request at 000000000001e440
> [38764.623677] IP: [<ffffffff814c8efe>] _spin_lock+0xe/0x30
> [38764.623677] PGD 12c7b4067 PUD 12c7b5067 PMD 0
> [38764.623677] Oops: 0002 [#2] SMP
> [38764.623677] last sysfs file: /sys/devices/virtual/block/ram9/stat
> [38764.623677] CPU 1
> or sometimes strangely affecting HP smart array, and causing it to
> disconnect it's raids (I don't understand how that's possible, but it
> doesn't happen with old openvz)
>
> Under the same load, classic 2.6.32-openvz kernels do just fine (although
> my personal feeling is that rhel6 is way more snappy under such a load).
>
> It usually takes less then few hours for rhel6 kernel to crash, although
> with lighter load it might take weeks or months.

I am very sad to hear this. Could you please file a bug to
bugzilla.openvz.org so our kernel guys will start working on that?

>
>
> Should we continue testing 'stable' branch, or maybe fixes are more likely
> to be expected in testing 042.x?
Well it depends. What we have in -testing branch is indeed testing, so
there can be more fixes but more bugs. Generally, if you have multiple
machines, I recommend to have a few (perhaps less important ones)
running rhel6-testing kernels, while having all the other ones at rhel6
(stable) branch.

The thing is, those -testing kernels are actually candidates for stable
repo, and as you can see some of them are then moved to stable (after we
do some internal testing to make sure there are no regressions etc).