Subject: Need help diagnosing OOM Killer report
Posted by rickb on Tue, 25 Sep 2007 20:10:48 GMT
View Forum Message <> Reply to Message

Hello Devs, I need your help.

Kernel: ovz028stab039.1
CPU: 2x quad core 2.33 xeon
mem: 32GB
disk: Emulex Corporation LP9000 Fibre Channel, local sata disks

When doing a really long read from the fiber channel card and write to the local disks, say 20+GB which takes 45 minutes or so, example when backing up a VE, the OOM is triggered:

http://208.77.99.251/oom.txt

Can you help me understand what is happening?

Thanks!
Rick

Subject: Re: Need help diagnosing OOM Killer report
Posted by rickb on Tue, 25 Sep 2007 20:23:37 GMT
View Forum Message <> Reply to Message

I asked almost the same question in may (different server/kernel).

http://forum.openvz.org/index.php?t=msg&goto=13373&

I believe this is causing the OOM killer to spawn as Vasily pointed it out last time:

Normal free:2728kB min:3756kB low:4692kB high:5632kB active:2324kB inactive:2624kB present:901120kB pages_scanned:7828 all_unreclaimable? yes

I notice this slab is rather enormous:

buffer_head : size 372822016 objsize 52

--

Any pointers or insight into what could be causing this or how I can avoid running out of normal_free would be great. To duplicate this problem I reported today, I was executing a long running tar of a many files.

Thanks!
Rick

Subject: Re: Need help diagnosing OOM Killer report
Posted by vaverin on Wed, 26 Sep 2007 12:17:13 GMT
View Forum Message <> Reply to Message

Rick,

could you please tell me what the kernel you have used?

OOM happens because disk activity eats some part of memory in normal zone and it becomes full.
whole normal zone is ~800mb and only buffer_head uses ~ 370Mb.
You can try to call sync in cycle (to force write to disk, it should reduce number of busy buffer_heads) -- but I'm not sure that it would be enough to prevent OOM-kill.

The simplest way to work around this issue -- increase size of normal zone.
In your kernel you have only 800 Mb used for normal zone.
Youcan switch to our enterprise kernels that have enabled 4Gb split option (at leaset for 2.6.9/rhel4 and 2.6.18/rhel5 kernels). 4Gb split patch increases size of normal zone up to 3.6 Gb.

Also you can switch to 64-bit kernels -- however it requires re-installation of hostOS on your node.
In general I would recommend to use 64-bit kernels now, even if you have only 32-bit VEs

thank you,
Vasily Averin

Subject: Re: Need help diagnosing OOM Killer report
Posted by rickb on Fri, 28 Sep 2007 18:16:49 GMT
View Forum Message <> Reply to Message

As suggested by Vasily via private chat, I patched 2.6.18 with the EL5 openvz kernel release, and it solved the problem. This allows 4GB of kernel memory space under 32bit rather then 800m.

Thanks!
Rick