Subject: Loadavg virtualisation problem (028stab047.1+fix)
Posted by Dariush Pietrzak on Mon, 12 Nov 2007 11:59:56 GMT
View Forum Message <> Reply to Message

Hello,
I encountered a ~6.0 load avg in idling guest, HN reports 0,0,0, cpuusage stats both from guest and from HN suggest that host is idling, but loadavg says 6.3,6.0,6.2 .
Temporary workaround: after 'vzctl chkpnt guest' and 'vzctl restore guest', guest started reporting correct values.
Guest is running only postgres database, both HN and guest are amd64 running on intel.

I've got dev machine with the same problem:
eyck@etchdev386:~/40m-ovz/work$ w
11:55:58 up 3 days, 2:41, 1 user, load average: 129.00, 129.00, 128.96
on HN:
codev64:/etc/vz/conf# w
11:56:20 up 3 days, 2:43, 7 users, load average: 0.00, 0.00, 0.00

the guest in question is really idling.
One thing that those two machines have in common, is that I played with adding/removing cpus from the guest (oh, and they are both amd64, although one guest is amd64 and another is i386).

--
Key fingerprint = 40D0 9FFB 9939 7320 8294 05E0 BCC7 02C4 75CC 50D9
Total Existance Failure

Subject: Re: Loadavg virtualisation problem (028stab047.1+fix)
Posted by dev on Sun, 18 Nov 2007 09:12:43 GMT
View Forum Message <> Reply to Message

Dariush,

1. what kernel do you use?
2. how do you decide that your machine is idling?
   can you please provide output of `ps axf` from HN?
3. what do you mean by "I played with adding/removing cpus from the guest"?
   vzctl set <VEID> --cpus <N> or something else?

Yep, I see a problem with counters in the code on --cpus decreasing...
will check...
http://bugzilla.openvz.org/show_bug.cgi?id=732

Thanks,
Kirill

Dariush Pietrzak wrote:
> Hello,
> I encountered a ~6.0 load avg in idling guest, HN reports 0,0,0, cpuusage
> stats both from guest and from HN suggest that host is idling, but loadavg
> says 6.3,6.0,6.2 .
> Temporary workaround: after 'vzctl chkpnt guest' and 'vzctl restore guest',
> guest started reporting correct values.
> Guest is running only postgres database, both HN and guest are amd64
> running on intel.
>
> I've got dev machine with the same problem:
> eyck@etchdev386:~/40m-ovz/work$ w
> 11:55:58 up 3 days, 2:41, 1 user, load average: 129.00, 129.00, 128.96
> on HN:
> codev64:/etc/vz/conf# w
> 11:56:20 up 3 days, 2:43, 7 users, load average: 0.00, 0.00, 0.00
>
> the guest in question is really idling.
> One thing that those two machines have in common, is that I played with
> adding/removing cpus from the guest (oh, and they are both amd64, although
> one guest is amd64 and another is i386).
>

Subject: Re: Loadavg virtualisation problem (028stab047.1+fix)
Posted by Dariush Pietrzak on Sun, 18 Nov 2007 10:53:08 GMT
View Forum Message <> Reply to Message

Kirill,
> 1. what kernel do you use?
 2.6.18 + ovz028stab047.1

> 2. how do you decide that your machine is idling?
>   can you please provide output of `ps axf` from HN?
 this sounds like philosophical question, howether - I created this machine
for compilation and load testing, so when I'm not loading it with anything,
it should be idling, and I confirm this with top/munin stats;)

> 3. what do you mean by "I played with adding/removing cpus from the guest"?
>   vzctl set <VEID> --cpus <N> or something else?
 Yes, I started with 8, and then did load testing with varying number of
cpus, this includes reducing number of cpus while the jobs are running
(which btw, results in 'top' and similiar tools failing with:
"top: failed /proc/stat read", strange that noone noticed this until now)

 The problem also goes away after I remove limits from number of cpus and
do some load testing, but this takes multiple hours (and in the meantime

guest that should be reporting load hovering in 0.2 ~ 1.0 range reported
load ~ 1500).
--
Key fingerprint = 40D0 9FFB 9939 7320 8294 05E0 BCC7 02C4 75CC 50D9
 Total Existance Failure

Subject: Re: Loadavg virtualisation problem (028stab047.1+fix)
Posted by dev on Sun, 18 Nov 2007 13:24:53 GMT
View Forum Message <> Reply to Message

Dariush,

I would be thankful if you find/provide some exact way
to reproduce it. If not - we will play a bit later and check it.
The issue is not major, it's most likely screw up of process
counters which doesn't affect anything but statistics.

Thanks,
Kirill
P.S. philosophical questions was due to the fact that
   top and similar tools don't show processes which took little
   CPU time and die, thus you can have 100% CPU load and
   top still doesn't show any processes consuming CPU.

Dariush Pietrzak wrote:
> Kirill,
>
>>1. what kernel do you use?
>
> 2.6.18 + ovz028stab047.1
>
>
>>2. how do you decide that your machine is idling?
>>   can you please provide output of `ps axf` from HN?
>
> this sounds like philosophical question, howether - I created this machine
> for compilation and load testing, so when I'm not loading it with anything,
> it should be idling, and I confirm this with top/munin stats;)
>
>
>>3. what do you mean by "I played with adding/removing cpus from the guest"?
>>   vzctl set <VEID> --cpus <N> or something else?
>
> Yes, I started with 8, and then did load testing with varying number of
> cpus, this includes reducing number of cpus while the jobs are running
> (which btw, results in 'top' and similiar tools failing with:
> "top: failed /proc/stat read", strange that noone noticed this until now)

>
> The problem also goes away after I remove limits from number of cpus and
> do some load testing, but this takes multiple hours (and in the meantime
> guest that should be reporting load hovering in 0.2 ~ 1.0 range reported
> load ~ 1500).

Subject: Re: Loadavg virtualisation problem (028stab047.1+fix)
Posted by Dariush Pietrzak on Sun, 18 Nov 2007 15:53:13 GMT
View Forum Message <> Reply to Message

Kirill,

> I would be thankful if you find/provide some exact way
> to reproduce it. If not - we will play a bit later and check it.
I just tried it like described below, and got result in ~1hour, I'll try tomorrow to find faster/more exact way of causing this, but right now what causes fairly consistently such result is this:
- I've got debian guest, with dev packages (kernel-package), and kernel sources sitting ready to compile:
- ssh to guest, export CONCURRENCY_LEVEL=6; fakeroot make-kpkg
- from HN, reduce CPUs to less then CONCURRENCY_LEVEL, for example to 1
when the make-kpkg finishes, what I have is guest that reports load_avg ~6:
eyck@etchdev386:~/40p-ovz/work$ w
15:39:52 up 4:47, 1 user, load average: 6.00, 6.01, 5.91
and host which correctly reports load ~0:
codev64:~# w
15:44:18 up 4:55, 2 users, load average: 0.00, 0.00, 0.24

I guess that what causes it, is having in guest more processes with runnable state then reduced virtual cpus available.

> The issue is not major, it's most likely screw up of process
> counters which doesn't affect anything but statistics.
yupp, it's almost cosmetic, but still, it's nicer to know that you're running code with little amount of warts, and also for people who monitor their guests from inside - this might result in unnecessary nightly alerts.

> Kirill
> P.S. philosophical questions was due to the fact that
> top and similar tools don't show processes which took little
> CPU time and die, thus you can have 100% CPU load and
> top still doesn't show any processes consuming CPU.
but then you should have this load visible on HN, and tools like sa should catch it (although I ain't got sa statistics on machine I detected this available).
--

Key fingerprint = 40D0 9FFB 9939 7320 8294 05E0 BCC7 02C4 75CC 50D9
 Total Existance Failure

Subject: Re: Loadavg virtualisation problem (028stab047.1+fix)
Posted by Dariush Pietrzak on Sun, 18 Nov 2007 16:18:46 GMT
View Forum Message <> Reply to Message

> I guess that what causes it, is having in guest more processes with runnable
> state then reduced virtual cpus available.
 Faster/simpler way:
 - go to guest, run 8x dd if=/dev/zero of=/dev/null
 - load should fairly quickly start creeping up to 8.0
 - on HN set cpus to 1
 - let it run, then stop all the dd's

 result:
Guest:
eyck@etchdev386:~/40p-ovz/work$ w
 16:13:44 up 5:21, 1 user, load average: 2.99, 2.94, 2.20
(and stays like this, 15-minute average even grows)
HN:
codev64:~# w
 16:14:05 up 5:25, 2 users, load average: 0.00, 1.76, 2.65

Playing with dd and set --cpus I also managed to cause load ~8 on HN while
guest reported only 2.0

.. I guess noone is doing things like this on production systems, and even
if so, running chkpnt/restore as precaution when performing such changes
is not out of the question.
--
Key fingerprint = 40D0 9FFB 9939 7320 8294 05E0 BCC7 02C4 75CC 50D9
 Total Existance Failure

Subject: Re: Loadavg virtualisation problem (028stab047.1+fix)
Posted by dev on Sun, 18 Nov 2007 16:36:18 GMT
View Forum Message <> Reply to Message

ok, Thanks a lot! I've added this info to the bug.

Thanks,
Kirill

Dariush Pietrzak wrote:
>>I guess that what causes it, is having in guest more processes with runnable

>>state then reduced virtual cpus available.
>
> Faster/simpler way:
> - go to guest, run 8x dd if=/dev/zero of=/dev/null
> - load should fairly quickly start creeping up to 8.0
> - on HN set cpus to 1
> - let it run, then stop all the dd's
>
> result:
> Guest:
> eyck@etchdev386:~/40p-ovz/work$ w
> 16:13:44 up 5:21, 1 user, load average: 2.99, 2.94, 2.20
> (and stays like this, 15-minute average even grows)
> HN:
> codev64:~# w
> 16:14:05 up 5:25, 2 users, load average: 0.00, 1.76, 2.65
>
> Playing with dd and set --cpus I also managed to cause load ~8 on HN while
> guest reported only 2.0
>
> .. I guess noone is doing things like this on production systems, and even
> if so, running chkpnt/restore as precaution when performing such changes
> is not out of the question.
