Subject: Limiting CPU for tasks on HN ?
Posted by geejay on Sat, 27 Sep 2008 10:09:43 GMT
View Forum Message <> Reply to Message

Hello,

It appears that CPU intensive tasks on the hardware node do kill completely the performance on the VEs.

E.g. Simply copying a big directory like a VE private directory on the hardware node, bringing CPU load to 20-25% or so, completely kills responsiveness of my VEs, making them basically unusuable.

Starting tasks with nice -9 doesnt help either.

Is there any way to limit the CPU usage of the Hardware node so that VEs can continue running ? Or can one guarantee minimum CPU time in VEs ?

As it stands I am unable to do essential maintenance tasks on the HN because VEs get effectively killed.

TIA for any input on the matter

I am running Centos 5
2.6.18-92.1.1.el5.028stab057.2

Geejay

Subject: Re: Limiting CPU for tasks on HN ?
Posted by maratrus on Sat, 27 Sep 2008 20:33:25 GMT
View Forum Message <> Reply to Message

Hello,

OpenVZ user's guide
http://download.openvz.org/doc/OpenVZ-Users-Guide.pdf
Managing Resources --> Managing CPU Share

Subject: Re: Limiting CPU for tasks on HN ?
Posted by geejay on Sun, 28 Sep 2008 06:24:27 GMT
View Forum Message <> Reply to Message

maratrus wrote on Sat, 27 September 2008 16:33Hello,

OpenVZ user's guide
http://download.openvz.org/doc/OpenVZ-Users-Guide.pdf
Managing Resources --> Managing CPU Share

Maratrus,

```
vzcpucheck  -v
VEID          CPUUNITS
------------------------
0             1000
102           83333
Current CPU utilization: 84333
Power of the node: 159866
```

there is only one VPS on the HN which has been allocated 80000 CPUUNITS

If I do a copy of several GB on the HN the VPS becomes unresponsive. The VPS is running a VoIP application, Asterisk.

SIP clients cannot initiate new calls while the copying on the HN happens. Typically the VPS would use maximum 2% CPU to initiate a call. So there should not be any problem with resources. Beancounters also show that everything is running fine. Also I can see that the VPS receives the SIP messages over the network, but Asterisk stalls and does not process the call. IMVHO the CPU scheduler does not work.

Something is definitely not working as suggested in the Openvz manual.

How can debug this problem ?

I should perhaps mention that the VPS was converted from a physical machine. On startup it throws the following error:

```
Sep 28 08:26:49 onl3 sysctl: error: unknown error 1 setting key 'kernel.sysrq'
Sep 28 08:26:49 onl3 sysctl: error: unknown error 1 setting key 'kernel.core_uses_pid'
Sep 28 08:26:49 online rc.sysinit: Configuring kernel parameters:  failed
```

However on the HN systcl shows no errors:

```
sysctl -p
net.ipv4.ip_forward = 1
kernel.shmmax = 4294967295
kernel.msgmax = 65536
kernel.msgmnb = 65536
net.ipv4.conf.all.rp_filter = 1
net.ipv4.conf.default.rp_filter = 1
net.ipv4.icmp_echo_ignore_broadcasts = 1
net.ipv4.conf.default.proxy_arp = 0
kernel.sysrq = 1
```

net.ipv4.conf.default.forwarding = 1
net.ipv4.conf.default.accept_source_route = 0
net.ipv4.conf.all.send_redirects = 0
kernel.shmall = 268435456
net.ipv4.conf.default.send_redirects = 1
kernel.core_uses_pid = 1
net.ipv4.tcp_syncookies = 1

cat /proc/user_beancounters
Version: 2.5

uid	resource	held	maxheld	barrier	limit	failcnt
102:	kmemsize	3265249	4588138	11055923	11377049	0
	lockedpages	0	0	3000	3000	0
	privvmpages	18487	34351	250000	250000	0
	shmpages	1293	1309	21504	21504	0
	dummy	0	0	0	0	0
	numproc	57	67	240	240	0
	physpages	9821	24879	0	2147483647	0
	vmguarpages	0	0	33792	2147483647	0
	oomguarpages	9821	24879	200000	2147483647	0
	numtcpsock	9	23	360	360	0
	numflock	4	10	188	206	0
	numpty	2	2	16	16	0
	numsiginfo	0	9	256	256	0
	tcpsndbuf	87204	229160	1720320	2703360	0
	tcprcvbuf	147456	326676	1720320	2703360	0
	othersockbuf	34240	174280	1126080	2097152	0
	dgramrcvbuf	0	32412	500000	500000	0
	numothersock	22	32	360	360	0
	dcachesize	0	0	3409920	3624960	0
	numfile	1389	1546	9312	9312	0
	dummy	0	0	0	0	0
	dummy	0	0	0	0	0
	dummy	0	0	0	0	0
	numiptent	26	26	128	128	0

Subject: Re: Limiting CPU for tasks on HN ?
Posted by maratrus on Mon, 29 Sep 2008 08:04:36 GMT
View Forum Message <> Reply to Message

Hello,

Quote:
If I do a copy of several GB on the HN the VPS becomes unresponsive.

Here you did an I/O intensive operation. Could you possible conduct an experiment which involves not I/O but CPU intensive task just to narrow down our assumptions?

Quote:
Also I can see that the VPS receives the SIP messages over the network, but Asterisk stalls and does not process the call.

Is it possible to find out the state of the problem task for example via top? And also cat /proc/$PID/wchan, where $PID is Asterisk's PID?

Does Asterisk became responsive when copying process on the HN is finished?

Quote:
sysctl: error: unknown error 1 setting key 'kernel.sysrq'

I think you shouldn't worry. That means that the process inside VE tried to tune sysctl which prohibited to do inside that VE.

Subject: Re: Limiting CPU for tasks on HN ?
Posted by kenjy on Mon, 29 Sep 2008 16:13:55 GMT
View Forum Message <> Reply to Message

OpenVz works with the kernel scheduler, you can set cpuunits for giving priority's like this:

root@windu:~# vzcpucheck -v
VEID CPUUNITS
----
0 1000
101 4000
102 2000
103 3012
Current CPU utilization: 13024
Power of the node: 221180

a highest value means more power.

Subject: Re: Limiting CPU for tasks on HN ?
Posted by geejay on Sun, 02 Nov 2008 10:35:58 GMT
View Forum Message <> Reply to Message

I have made some further tests on the issue with ionice.

E.g. if I untar a file in the HN with
ionice -c3 -n7 tar -xzvf 122.tar.gz

Ionice with lowest possible priority in idle class.

Then also Asterisk becomes unresponsive inside the VM - although the untar on the HN becomes idle when Asterisks tries to start its task. As soon as the untar is finished Asterisk responds again.

It appears that it is impossible to limit IO related tasks on the HN so that Asterisk tasks in the VM does not get disturbed.

It may be the case that this is related to the timing driver that Asterisk uses. The timing driver ztdummy is installed on the HN and imported into the VM.

However I am stuck now, I dont know how to investigate this further.

In any case the impression remains that there is some bug or design error in OpenVZ, this kind of behaviour is clearly not desired. It effectively prevents one from doing any sort of heavier file tasks on the HN.

Geejay

Subject: Re: Limiting CPU for tasks on HN ?
Posted by seanfulton on Sun, 02 Nov 2008 19:05:26 GMT
View Forum Message <> Reply to Message

I am experiencing a similar problem. We moved our mail server from a physical machine with 2 G of RAM and 2 G of SWAP into a virtual machine with essentially the same resources.

Nightly backup causes the mail server activity to drop to nothing, causing back-ups and major issues. I have spent about a month trouble-shooting this and verified that there is nothing within the VE that is causing the hang up.

When I turn off the nightly backup, no problems. When I run one, everything hangs.

Two things I have noticed which are strange: there are no errors inside the VE, just more and more sendmail processes launching, but the old ones don't close.

Second, it is possible to allocate more RAM to a VE than the system has, causing the HN to crash. I opened up all of the memory settings for the VE until the HN crashed. Had to back off the settings to keep it up. I know when the machine was running on a physical machine, it would kill off processes as the system ran out of ram (and swap). This activity happens within the VE itself. But if you accidentally tell the VE it has 8 G available but it only has 6G, it will cause a kernel panic the HN once it exhausts the 6 G.

BTW: Nightly backup is vzdump to an NFS mount. THis is how I am set (2202 is the sendmail VE):

```
[root@gcn16 ~]# vzcpucheck -v
VEID          CPUUNITS
------------------------
0             1000
2202            100000
6005            1000
Current CPU utilization: 102000
Power of the node: 232811
[root@gcn16 ~]#
```

sean

Subject: Re: Limiting CPU for tasks on HN ?
Posted by geejay on Mon, 03 Nov 2008 08:35:13 GMT
View Forum Message <> Reply to Message

seanfulton wrote on Sun, 02 November 2008 14:05I am experiencing a similar problem.
.....
BTW: Nightly backup is vzdump to an NFS mount. THis is how I am set (2202 is the sendmail VE):

Sean, are you running also Centos 5 on the HN ?

Subject: Re: Limiting CPU for tasks on HN ?
Posted by ulver on Mon, 03 Nov 2008 08:40:49 GMT
View Forum Message <> Reply to Message

I have exactly the same problem, I tried to explain it in another topic (http://forum.openvz.org/index.php?t=msg&th=6884&star t=0&) but I didn't have any answer (my english is very bad and it's difficult for me to explain the problem correctly so I think it wasn't very clear).

When I make some I/O disk operations on the HN (backups with rsync, du, rm...), all the VM and the HN is geeting more and more slower... and a simple backup takes very long time. I didn't have the same problem with Vserver.

Subject: Re: Limiting CPU for tasks on HN ?
Posted by seanfulton on Mon, 03 Nov 2008 11:43:29 GMT
View Forum Message <> Reply to Message

Yes, as a matter of fact, I am.

Subject: Re: Limiting CPU for tasks on HN ?
Posted by ulver on Thu, 06 Nov 2008 10:17:52 GMT
View Forum Message <> Reply to Message

It seems that "jmctavish" in this topic (http://forum.openvz.org/index.php?t=msg&th=6868&start=0& have the same problem with IO performance.

Maratrus told him to upgrade his kernel, but we don't know if it worked.
In my case, I have the stable kernel release available in :
deb http://download.openvz.org/debian-systs etch openvz
(2.6.18-12-fza-amd64)

Subject: Re: Limiting CPU for tasks on HN ?
Posted by geejay on Thu, 06 Nov 2008 10:53:19 GMT
View Forum Message <> Reply to Message

On other machines we have also OpenVZ on Debian Etch and there I didnt notice such I/O problems - but I cannot test there properly because there are crucial tasks on that machine.

We have to setup a test machine to investigate the problem thoroughly.

Subject: Re: Limiting CPU for tasks on HN ?
Posted by seanfulton on Thu, 06 Nov 2008 16:19:17 GMT
View Forum Message <> Reply to Message

We have four machines running OpenVZ, various kernels. The one in question is running:

2.6.18-53.1.19.el5.028stab053.14PAE

I can't say for sure if this happens on our other machines because those are mostly web sites and we run the backups at night when traffic is low, so it's possible the problem just isn't affecting anything. The mail server is obvious because that tends to churn 24/7 so whenever I run my backups (using vzdump) everything hangs and backs up, causing it to eventually crash in the AM.

One thought I did have though is that even though the crash is during a period of high IO activity, the sendmail processes hang, stack up and eventually lock up the VE. When the backup completes, the sendmail activity doesn't seem to go back to normal--it seems to be permanently locked and just keeps spawning new processes until it crashes (usually at about 1800 processes).

This weekend I will upgrade the machine to the latest kernel and see if that makes a difference. I will report back.

sean

Subject: Re: Limiting CPU for tasks on HN ?
Posted by seanfulton on Mon, 10 Nov 2008 16:06:48 GMT
View Forum Message <> Reply to Message

seanfulton wrote on Thu, 06 November 2008 11:19We have four machines running OpenVZ, various kernels. The one in question is running:

2.6.18-53.1.19.el5.028stab053.14PAE

I can't say for sure if this happens on our other machines because those are mostly web sites and we run the backups at night when traffic is low, so it's possible the problem just isn't affecting anything. The mail server is obvious because that tends to churn 24/7 so whenever I run my backups (using vzdump) everything hangs and backs up, causing it to eventually crash in the AM.

One thought I did have though is that even though the crash is during a period of high IO activity, the sendmail processes hang, stack up and eventually lock up the VE. When the backup completes, the sendmail activity doesn't seem to go back to normal--it seems to be permanently locked and just keeps spawning new processes until it crashes (usually at about 1800 processes).

This weekend I will upgrade the machine to the latest kernel and see if that makes a difference. I will report back.

sean

Update: I upgraded to the 57.2 kernel and am having the same problems.

I was not able to update to the 59.3 kernel because there is a problem initializing the quota in that kernel.

Any ideas?

sean