Subject: mem limit does not work... ?
Posted by disaster on Sat, 18 Feb 2006 10:51:29 GMT
View Forum Message <> Reply to Message

Hello!

I've set a RAM Limit to 128MB - but the VPS can consume 2GB without any problems... that can't be right or?

I also have the following messages in syslog:
Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 262218 h 218587, res unused_privvmpages ub 4
Uncharging too much 2 h 0, res unused_privvmpages ub 4

Anybody here who has an idea?

Stefan

cat /proc/user_beancounters
Version: 2.5

uid	resource	held	maxheld	barrier	limit	failcnt
4:	kmemsize	434569	1484885	2752512	2936012	0
	lockedpages	0	0	32	32	0
	privvmpages	3910	263944	49152	53575	16
	shmpages	640	656	8192	8192	0
	dummy	0	0	0	0	0
	numproc	7	10	65	65	0
	physpages	944	263305	134217728	134217728	0
	vmguarpages	0	0	134217728	134217728	0
	oomguarpages	944	263305	134217728	134217728	0
	numtcpsock	2	2	80	80	0
	numflock	1	2	100	110	0
	numpty	1	1	16	16	0
	numsiginfo	0	6	256	256	0
	tcpsndbuf	2224	8896	319488	524288	0
	tcprcvbuf	0	8544	319488	524288	0
	othersockbuf	2224	7200	132096	336896	0
	dgramrcvbuf	0	4272	132096	132096	0
	numothersock	1	6	80	80	0
	dcachesize	91840	95858	1048576	1097728	0
	numfile	86	111	2048	2048	0
	dummy	0	0	0	0	0
	dummy	0	0	0	0	0
	dummy	0	0	0	0	0
	numiptent	4	4	128	128	0

Subject: Re: mem limit does not work... ?
Posted by dev on Sun, 19 Feb 2006 07:27:28 GMT
View Forum Message <> Reply to Message

All the parameters which have "pages" in their names are configured in pages...
So you set 4096*128Mb limits... it is much more than 2GB

We will add simple ability to specify UBC parameters in any units in vzctl soon, so e.g. you will be able to do --vmguarpages 128m.

Subject: Re: mem limit does not work... ?
Posted by disaster on Sun, 19 Feb 2006 09:11:53 GMT
View Forum Message <> Reply to Message

Thanks for your reply - so if i want to set a limit to 128M i have now tried - the following:
128M*1024*1024 => 134217728 bytes / 4096 => 32768

and then i used:
vzctl set 4 --privvmpages 32768:32768 --oomguarpages 32768:32768 --physpages 32768:32768 --vmguarpages 32768:32768 --save

but it dos not work... i can still allocate 2GB...

Another question:
Is there a listing for the meaning of privvmpages, vmguarpages, numpty, numproc etc. ?

Stefan

Subject: Re: mem limit does not work... ?
Posted by disaster on Sun, 19 Feb 2006 09:46:27 GMT
View Forum Message <> Reply to Message

OK - i found some strange things...

1. i can still allocate 1GB and more
2. a cat cat /proc/user_beancounters shows me failures at privvmpages failcnt=4 - but the script can work and is not killed
3. on the host machine i see in syslog:

Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 1 h 0, res unused_privvmpages ub 4

Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 1 h 0, res unused_privvmpages ub 4
Uncharging too much 1 h 0, res unused_privvmpages ub 4
Fatal resource shortage: kmemsize, UB 4.
Uncharging too much 65610 h 54827, res unused_privvmpages ub 4
Uncharging too much 3 h 0, res unused_privvmpages ub 4

What does that mean?

4. if i set kmemsize to 655360 the script can only allocate 128MB

Any idea?

Subject: Re: mem limit does not work... ?
Posted by dev on Sun, 19 Feb 2006 09:57:39 GMT
View Forum Message <> Reply to Message

1. can you please provide /proc/user_beancounters and a script?...

2. failct=4 and messages like:
Fatal resource shortage: kmemsize, UB 4.
mean that you hit kmemsize limit 4 times.

Applications are not killed in these cases. Usually syscalls simply return ENOMEM and most applications handle errors gracefully.

3. messages like:
Uncharging too much 1 h 0, res unused_privvmpages ub 4
mean that there is some discrepancy in privvmpages accounting. This is why I would be happy to your script which reproduces this problem.

4. kmemsize has nothing to do with app memory.

So to answer your questions quickly it would be helpfull to get your script and exact examples/steps you used and wonder about their results.

Subject: Re: mem limit does not work... ?
Posted by disaster on Sun, 19 Feb 2006 10:26:10 GMT
View Forum Message <> Reply to Message

Quote:

1. can you please provide /proc/user_beancounters and a script?...

No Problem - here are the details:

user_beancounters before script:

```
cat /proc/user_beancounters
Version: 2.5
      uid  resource          held    maxheld    barrier      limit    failcnt
        4: kmemsize        434432     477994   13107200   13107200          0
           lockedpages          0          0         32         32          0
           privvmpages       3915       5868      32768      32768          0
           shmpages           640        656       8192       8192          0
           dummy                0          0          0          0          0
           numproc              7          8         65         65          0
           physpages          945        945      32768      32768          0
           vmguarpages          0          0      32768      32768          0
           oomguarpages       945        945      32768      32768          0
           numtcpsock           2          2         80         80          0
           numflock             1          2        100        110          0
           numpty               1          1         16         16          0
           numsiginfo           0          1        256        256          0
           tcpsndbuf         2224       8896     319488     524288          0
           tcprcvbuf            0       4272     319488     524288          0
           othersockbuf      2224       7200     132096     336896          0
           dgramrcvbuf          0       4272     132096     132096          0
           numothersock         1          5         80         80          0
           dcachesize       91840      94136    1048576    1097728          0
           numfile             86         90       2048       2048          0
           dummy                0          0          0          0          0
           dummy                0          0          0          0          0
           dummy                0          0          0          0          0
           numiptent            4          4        128        128          0
```

My Script - you don't like... and this time with top output

```
perl -e '$a="a";for(;;) {$a.=$a; print "allocated ".int(length($a)/1024/1024)."MB\n";}'
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
```

allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 0MB
allocated 1MB
allocated 2MB
allocated 4MB
allocated 8MB
allocated 16MB
allocated 32MB
allocated 64MB
allocated 128MB
allocated 256MB
allocated 512MB
allocated 1024MB
Out of memory!

Top Output

Mem: 2071904k total, 1067136k used, 1004768k free, 10272k buffers
Swap: 1437776k total, 0k used, 1437776k free, 46148k cached

PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ Command

4544 root 25 0 1027m 878m 2576 R 94.0 43.4 0:09.74 perl

As you can see, there is a total of 1GB RAM and perl eats 43% which is about 800MB the res 200MB are used from normal system.

user_beancounters after script start

cat /proc/user_beancounters
Version: 2.5

uid	resource	held	maxheld	barrier	limit	failcnt
4:	kmemsize	434432	1485574	13107200	13107200	0
	lockedpages	0	0	32	32	0
	privvmpages	1621	263957	32768	32768	10
	shmpages	640	656	8192	8192	0
	dummy	0	0	0	0	0
	numproc	7	8	65	65	0
	physpages	946	263318	32768	32768	0

vmguarpages	0	0	32768	32768	0
oomguarpages	946	263318	32768	32768	0
numtcpsock	2	2	80	80	0
numflock	1	2	100	110	0
numpty	1	1	16	16	0
numsiginfo	0	1	256	256	0
tcpsndbuf	2224	8896	319488	524288	0
tcprcvbuf	0	4272	319488	524288	0
othersockbuf	2224	7200	132096	336896	0
dgramrcvbuf	0	4272	132096	132096	0
numothersock	1	5	80	80	0
dcachesize	91840	94136	1048576	1097728	0
numfile	86	90	2048	2048	0
dummy	0	0	0	0	0
dummy	0	0	0	0	0
dummy	0	0	0	0	0
numiptent	4	4	128	128	0

Quote:
2. failct=4 and messages like:
Fatal resource shortage: kmemsize, UB 4.
mean that you hit kmemsize limit 4 times.

Sorry but failcnt is at privvmpages not at kmemsize (at kmemsize only in my example before...).

Quote:
3. messages like:
Uncharging too much 1 h 0, res unused_privvmpages ub 4
mean that there is some discrepancy in privvmpages accounting.
This is why I would be happy to your script which reproduces this problem.

No problem - you have the small one line script above...

I would be happy if you can help me - if you need anything - write it down - i'll do that.

Thanks!

Subject: Re: mem limit does not work... ?
Posted by disaster on Sun, 19 Feb 2006 13:07:28 GMT
View Forum Message <> Reply to Message

I've now tested Kernel 2.6.15 with the beta Patch - there it WORKS!!! I'm not able to allocate more that 115MB with perl - but the fairscheduler does not work

So i think we must wait for fairscheduler 2.6.15 or for a patch for 2.6.8 ...

Stefan

Subject: Re: mem limit does not work... ?
Posted by dev on Sun, 19 Feb 2006 17:56:25 GMT
View Forum Message <> Reply to Message

Don't worry, I will check it soon and will return back to you.

Subject: Re: mem limit does not work... ?
Posted by dev on Mon, 20 Feb 2006 10:23:38 GMT
View Forum Message <> Reply to Message

I filled in bug #104 (http://bugzilla.openvz.org/show_bug.cgi?id=104)
The patch is attached inside.
New kernel will be released soon.

Thanks a lot for reporting this!

Subject: Re: mem limit does not work... ?
Posted by disaster on Mon, 20 Feb 2006 10:29:42 GMT
View Forum Message <> Reply to Message

Thanks a lot! I will patch the kernel and will try it again. Also i will test some other things the next days. If i find some other things - i'll report this.

Subject: Re: mem limit does not work... ?
Posted by atomic on Mon, 20 Feb 2006 10:45:54 GMT
View Forum Message <> Reply to Message

Thanks for solving this bug! The patch works fine on my system.
