
Subject: High Load Issues

Posted by [aeterna](#) on Fri, 04 May 2007 11:56:13 GMT

[View Forum Message](#) <> [Reply to Message](#)

I have been experiencing high load on my HN. Sometimes the load get's as high as 100+. I've narrowed it down to improper configuration files for our VE's however I was hoping someone could just confirm my suspicions. I just started reading up on everything the wiki has but its still alot of information that im catching up on..

Here is an example of what one of our VE's UBC parameters look like:

```
# UBC parameters (in form of barrier:limit)
# Primary parameters
AVNUMPROC="80"
NUMPROC="385"
NUMTCPSOCK="160"
NUMOTHERSOCK="2147483647"
VMGUARPAGES="12288:2147483647"
# Secondary parameters
KMEMSIZE="16777216:16777216"
TCPSNDBUF="4880000:8000000"
TCPRCVBUF="4880000:8000000"
OTHERSOCKBUF="132096:336896"
DGRAMRCVBUF="1058384:1058384"
OOMGUARPAGES="16144:2147483647"
# Auxiliary parameters
LOCKEDPAGES="32:32"
SHMPAGES="8192"
PRIVVMPAGES="351072:422978"
NUMFILE="14280"
NUMFLOCK="100:110"
NUMPTY="36"
NUMSIGINFO="256"
DCACHESIZE="5916672:6163200"
PHYSPAGES="0:2147483647"
NUMIPTENT="2147483647"
```

So if i take $12288 * 4096 = 50331648$ bytes = 48 MB = This is my guaranteed memory limit?

while my burstable is something like 1.3GB, correct?

Subject: Re: High Load Issues

Posted by [dev](#) on Fri, 04 May 2007 12:49:05 GMT

[View Forum Message](#) <> [Reply to Message](#)

by "load" I guess you mean "load average" reported by top and uptime commands, right?

Load average can be due to a number of reasons:

- you have set cpulimits to VE
- too much disk I/O (many tasks in D state waiting for I/O)
- some bug/deadlock (some task is always in D state)
- many CPU-bound applications running (tasks in R state)

Load average itself is a sum of number of tasks in R and D states. So you need to investigate task states via:

```
# ps axf
```

```
# ps axf -o pid,wchan,comm
```

```
# Alt-SysRq-T
```

Subject: Re: High Load Issues

Posted by [aeterna](#) on Fri, 04 May 2007 13:06:12 GMT

[View Forum Message](#) <> [Reply to Message](#)

yea, im referring to the what im seeing in my top commands. But i think whoever initially set this up didn't set the basic configuration for our vps's correctly. Initially we had cpulimits set to a certain value however were instructed to set them all to 0 or to unlimited, which didnt make any sense to me.

Was I correct with my equation about the VMGUARPAGES/PRIVVPAGES variable?

btw thanks for the reply and help

Subject: Re: High Load Issues

Posted by [rickb](#) on Fri, 04 May 2007 13:43:15 GMT

[View Forum Message](#) <> [Reply to Message](#)

High load on the HN means many procs in the runqueue. It does not necessarily mean 0% idle cpu (as it usually does in a non HN/vps server). many procs and idle cpu occurs when VEs have a low cpu limit and they are sustaining that limit. So, does your HN have idle cpu while having 100 load?

Subject: Re: High Load Issues

Posted by [dev](#) on Fri, 04 May 2007 15:19:59 GMT

[View Forum Message](#) <> [Reply to Message](#)

not sure what you mean by "which didnt make any sense to me".
do you have cpulimits or not? plz do the following:

cat /proc/fairsched2
and post it here...

yes, your math is correct.

you can do:

vzctl set VEID --vmguarpages 48m --save
instead of calculating pages.

Subject: Re: High Load Issues
Posted by [aeterna](#) on Fri, 04 May 2007 19:44:51 GMT
[View Forum Message](#) <> [Reply to Message](#)

cat /proc/fairsched
Version: 2.6 debug

veid	id	parent	weight	rate	tasks	run	cpus	flg	ready	start_tag	value	
0	0	0	14	0	225	226	1	..	1	94611247807425	0	
0	0	0	14	0	225	226	1	..	1	94611247807425	0	
0	90	0	31	4096	40	0	0	L.	0	94613261814411	17364433920	
4373176203264498	90	0	90	31	4096	40	0	0	L.	0	94613261814411	17364433920
4373176203264498	90	1	0	31	4096	40	0	0	L.	0	94613261814411	17364433920
4373176203264498	0	1	0	500	0	33	0	0	..	0	94613389381069	0
0	1	0	1	500	0	33	0	0	..	0	94613389381069	0
0	1	1	0	500	0	33	0	0	..	0	94613389381069	0
0	31	0	31	4096	57	0	0	L.	0	94589523243940	17364433920	
4372836721617697	31	0	31	31	4096	57	0	0	L.	0	94589523243940	17364433920
4372836721617697	31	1	0	31	4096	57	0	0	L.	0	94589523243940	17364433920
4372836721617697	0	83	0	31	4096	45	0	0	L.	0	94616983438502	17364433920
4372997905031502	83	0	83	31	4096	45	0	0	L.	0	94616983438502	17364433920
4372997905031502	83	1	0	31	4096	45	0	0	L.	0	94616983438502	17364433920
4372997905031502	0	88	0	31	4096	60	0	0	L.	0	94583732263958	17364433920
4373090581760476												

88	0	88	31	4096	60	0	0	L.	0	94583732263958	17364433920
4373090581760476											
88	1	0	31	4096	60	0	0	L.	0	94583732263958	17364433920
4373090581760476											
0	91	0	15	4096	42	0	0	L.	0	94580675582920	17364433920
4373154798021511											
91	0	91	15	4096	42	0	0	L.	0	94580675582920	17364433920
4373154798021511											
91	1	0	15	4096	42	0	0	L.	0	94580675582920	17364433920
4373154798021511											
0	94	0	31	4096	51	0	0	L.	0	94612780609535	17364433920
4373057796476552											
94	0	94	31	4096	51	0	0	L.	0	94612780609535	17364433920
4373057796476552											
94	1	0	31	4096	51	0	0	L.	0	94612780609535	17364433920
4373057796476552											
0	99	0	31	4096	42	0	0	L.	0	94613468898127	17364433920
4373211672662200											
99	0	99	31	4096	42	0	0	L.	0	94613468898127	17364433920
4373211672662200											
99	1	0	31	4096	42	0	0	L.	0	94613468898127	17364433920
4373211672662200											
0	100	0	31	4096	46	0	0	L.	0	94610065335620	17364433920
4372522768509284											
100	0	100	31	4096	46	0	0	L.	0	94610065335620	
17364433920	4372522768509284										
100	1	0	31	4096	46	0	0	L.	0	94610065335620	17364433920
4372522768509284											
0	103	0	31	4096	42	0	0	L.	0	94610619084712	17364433920
4373317617056216											
103	0	103	31	4096	42	0	0	L.	0	94610619084712	
17364433920	4373317617056216										
103	1	0	31	4096	42	0	0	L.	0	94610619084712	17364433920
4373317617056216											
0	104	0	15	4096	47	0	0	L.	0	94591515911221	17364433920
6005840073060709											
104	0	104	15	4096	47	0	0	L.	0	94591515911221	
17364433920	6005840073060709										
104	1	0	15	4096	47	0	0	L.	0	94591515911221	17364433920
6005840073060709											
0	169	0	15	4096	70	0	0	L.	0	94618630219252	17364433920
4373523471763362											
169	0	169	15	4096	70	0	0	L.	0	94618630219252	
17364433920	4373523471763362										
169	1	0	15	4096	70	0	0	L.	0	94618630219252	17364433920
4373523471763362											
0	179	0	15	4096	42	0	0	L.	0	94600751818083	17364433920
4373576791826538											

179	0	179	15	4096	42	0	0	L.	0	94600751818083	
17364433920	4373576791826538										
179	1	0	15	4096	42	0	0	L.	0	94600751818083	17364433920
4373576791826538											
0	479	0	31	4096	69	0	0	L.	0	94581043769286	17364433920
4372214340662038											
479	0	479	31	4096	69	0	0	L.	0	94581043769286	
17364433920	4372214340662038										
479	1	0	31	4096	69	0	0	L.	0	94581043769286	17364433920
4372214340662038											
0	199	0	31	4096	42	0	0	L.	0	94579551579246	17364433920
4373682218820909											
199	0	199	31	4096	42	0	0	L.	0	94579551579246	
17364433920	4373682218820909										
199	1	0	31	4096	42	0	0	L.	0	94579551579246	17364433920
4373682218820909											
0	480	0	31	4096	68	0	0	L.	0	94606770318073	17364433920
4925320687894040											
480	0	480	31	4096	68	0	0	L.	0	94606770318073	
17364433920	4925320687894040										
480	1	0	31	4096	68	0	0	L.	0	94606770318073	17364433920
4925320687894040											
0	211	0	15	4096	45	0	0	L.	0	94613455871164	17364433920
4373716229002812											
211	0	211	15	4096	45	0	0	L.	0	94613455871164	
17364433920	4373716229002812										
211	1	0	15	4096	45	0	0	L.	0	94613455871164	17364433920
4373716229002812											
0	92	0	31	4096	48	0	0	L.	0	94617287409219	17364433920
4373258475828748											
92	0	92	31	4096	48	0	0	L.	0	94617287409219	17364433920
4373258475828748											
92	1	0	31	4096	48	0	0	L.	0	94617287409219	17364433920
4373258475828748											
0	202	0	31	4096	48	0	0	L.	0	94610614577033	17364433920
8191263418904807											
202	0	202	31	4096	48	0	0	L.	0	94610614577033	
17364433920	8191263418904807										
202	1	0	31	4096	48	0	0	L.	0	94610614577033	17364433920
8191263418904807											
0	89	0	15	4096	49	0	0	L.	0	94613011455904	17364433920
4377486407989553											
89	0	89	15	4096	49	0	0	L.	0	94613011455904	17364433920
4377486407989553											
89	1	0	15	4096	49	0	0	L.	0	94613011455904	17364433920
4377486407989553											
0	93	0	31	4096	49	0	0	L.	0	94619171211298	17364433920
0											

93	0	93	31	4096	49	0	0	L.	0	94619171211298	17364433920
	0										
93	1	0	31	4096	49	0	0	L.	0	94619171211298	17364433920
	0										
0	87	0	7	4096	45	0	0	L.	0	94607604219838	17364433920
7564054593336869											
87	0	87	7	4096	45	0	0	L.	0	94607604219838	17364433920
7564054593336869											
87	1	0	7	4096	45	0	0	L.	0	94607604219838	17364433920
7564054593336869											
0	451	0	31	4096	63	0	0	L.	0	94602019621982	17364433920
	0										
451	0	451	31	4096	63	0	0	L.	0	94602019621982	
17364433920			0								
451	1	0	31	4096	63	0	0	L.	0	94602019621982	17364433920
	0										
0	86	0	31	4096	72	0	0	L.	0	94609539513109	17364433920
	0										
86	0	86	31	4096	72	0	0	L.	0	94609539513109	17364433920
	0										
86	1	0	31	4096	72	0	0	L.	0	94609539513109	17364433920
	0										
0	175	0	31	4096	62	0	0	L.	0	94589211757924	17364433920
	0										
175	0	175	31	4096	62	0	0	L.	0	94589211757924	
17364433920			0								
175	1	0	31	4096	62	0	0	L.	0	94589211757924	17364433920
	0										
0	35	0	15	4096	62	0	0	L.	0	94609510117362	17364433920
	0										
35	0	35	15	4096	62	0	0	L.	0	94609510117362	17364433920
	0										
35	1	0	15	4096	62	0	0	L.	0	94609510117362	17364433920
	0										
0	78	0	15	4096	57	0	0	L.	0	94614092806234	17364433920
	0										
78	0	78	15	4096	57	0	0	L.	0	94614092806234	17364433920
	0										
78	1	0	15	4096	57	0	0	L.	0	94614092806234	17364433920
	0										
0	192	0	7	4096	41	0	0	L.	0	94603905076716	17364433920
	0										
192	0	192	7	4096	41	0	0	L.	0	94603905076716	
17364433920			0								
192	1	0	7	4096	41	0	0	L.	0	94603905076716	17364433920
	0										
0	17	0	4	4096	83	0	0	L.	0	94608126511294	17364433920
	0										

```
17      0      17   4 4096  83   0   0 L.   0   94608126511294   17364433920
      0
17      1      0   4 4096  83   0   0 L.   0   94608126511294   17364433920
      0
```

All my vps have their cpulimits set to 0.

Like I said I just started reading about this stuff this week. CPULIMITS is the % of time that each individual VPS has with the nodes cpu..

My statement about 'it not making sense' was that if you're experiencing high load on a node how would setting everything to a 0 or unlimited help reduce the load?

Subject: Re: High Load Issues

Posted by [dev](#) on Sun, 06 May 2007 16:58:20 GMT

[View Forum Message](#) <> [Reply to Message](#)

ok. you don't have cpulimited VEs.

No you need to dump:

```
# ps axf -o pid,wchan,comm
```

to see which processes are in D state and where.

Subject: Re: High Load Issues

Posted by [rickb](#) on Sun, 06 May 2007 18:07:28 GMT

[View Forum Message](#) <> [Reply to Message](#)

I think its worth noting--

As Dev pointed out, you do not have any cpulimits set on your VEs. This is generally bad especially if each VE has its own admin; ie you want to ensure fairness rather than optimal overall efficiency. When you add cpulimits to your VEs, the load will not decrease, and in fact it may increase. This is because a few of your VEs will reach their cpulimit and their runqueues will expand, which means more processes want to run as seen by your HN.

However, the end result is your cpuunits will kick into gear and each environment will have its amount of cpuunits cputime/equivalent MHz to count on. Each of your VEs will feel more interactive when cpuunits are well enforced, and cpulimit will protect the HN from reaching 0% idle time very often, which makes things even more interactive.

Be sure that your cpuunits are not oversubscribed. See my message #8620 on this page to read more about cpuunits/cpulimit and how they relate to each other.

<http://forum.openvz.org/index.php?&t=msg&th=1551>

The bottom line is that the load as seen by the HN is not always an accurate factor in determining if your server is overcommitted, underpowered, etc (you may have plenty of idle cpu time with a load of 100!). With regard to strictly your application load, it may and probably is overcommitted; however if you are not the admin of your VEs, like in the case you are selling them as a service, you are fulfilling your responsibility in guaranteeing them a time slice of the cpu / equivalent CPU speed in MHz and the HN load is not something you can control.

Hope this isn't too wordy!