Subject: Memory Management again
Posted by atomic on Sun, 19 Feb 2006 10:34:17 GMT
View Forum Message <> Reply to Message

Hi,

i read many postings about memory restrictions and guaranteed memory sizes, so i figured out some
values for standard memory sizes.

VMGUARPAGES="32768:2147483647" // 128MB Guaranteed
VMGUARPAGES="65536:2147483647" // 256MB Guaranteed
VMGUARPAGES="131072:2147483647" // 512MB Guaranteed

PRIVVMPAGES="64000:65536" // 256MB Burst
PRIVVMPAGES="128000:131072" // 512MB Burst
PRIVVMPAGES="256000:262140" // 1024MB Burst

Calculation:
VMGUARPAGES and PRIVVMPAGES values are given in memory pages.
One memory page holds 4096Bits or 4Kbyte on a x86-System.
65536 Pages * 4Kbyte = 262144Kbyte = 256Mbyte

Where "guaranteed" means the absolute minimum of available RAM, even if the node is heavily loaded,
and "burst" means the limit of allocatable RAM if there is enough free.

Here is a beancounter output of a VPS with 256Mbyte guaranteed and 1024MByte burst memory:

Version: 2.5

uid	resource	held	maxheld	barrier	limit	failcnt
3:	kmemsize	3532503	4198588	16979722	18677694	0
	lockedpages	0	0	829	829	0
	privvmpages	147040	149045	256000	262140	0
	shmpages	11417	11433	31096	31096	0
	dummy	0	0	0	0	0
	numproc	100	117	800	800	0
	physpages	32964	34396	0	2147483647	0
	vmguarpages	0	0	65535	2147483647	0
	oomguarpages	32964	34396	310967	2147483647	0
	numtcpsock	24	25	800	800	0
	numflock	7	10	662	728	0
	numpty	1	1	80	80	0
	numsiginfo	0	2	1024	1024	0
	tcpsndbuf	4456	16744	2383107	5659907	0

tcprcvbuf	1204	9576	2383107	5659907	0
othersockbuf	20896	29268	1191553	4468353	0
dgramrcvbuf	0	2228	1191553	1191553	0
numothersock	24	30	800	800	0
dcachesize	322624	332280	3704295	3815424	0
numfile	719	777	6624	6624	0
dummy	0	0	0	0	0
dummy	0	0	0	0	0
dummy	0	0	0	0	0
numiptent	10	10	200	200	0

Are these values correct and expedient? What does the value "2147483647" at VMGUARPAGES mean?
2147483647 memory pages are roughly 8192 Gigabytes, is this the maximal allocatable memory size of OpenVZ?

Thanks.

Subject: Re: Memory Management again
Posted by Sten on Fri, 14 Dec 2007 14:50:25 GMT
View Forum Message <> Reply to Message

„Guaranteed“ (barrier) means any application in the VE can use the memory, „burst“ (limit) is the very maximum value of memory taken; memory between barrier and limit can be used by applications running under root only.

2147483647 is the maximal value of 32-bit unsigned integer, so if you’re running 32-bit system, it’s the maximum allocable memory of the OpenVZ.

Subject: Re: Memory Management again
Posted by ledude1 on Thu, 03 Jan 2008 03:19:33 GMT
View Forum Message <> Reply to Message

Atomic and Sten.

This is awesome example on how to calculate the memory and memory burst. I've been trying my dearest to understand exactly what this one means reading the manual and for some reason this one answer my question.

But I have one more question though and perhaps someone can help me understand it.

PRIVVMPAGES="256000:262140" // 1024MB Burst

I know it is for the memory burst but help me understand again how I do the calculation?
I understand that the 256000 is where you get the 1024 Mb burst but what's 262140? What's that mean on the bean counter by limit? Is it basically means that you can have range or memory burst from 1024 to 1048? Please help me understand it? I've read the manual and it makes no sense to me. Also how do I get the 262140?

Many thanks for all of your help and pardon me being a nooby.

Subject: Re: Memory Management again
Posted by vaverin on Thu, 03 Jan 2008 07:30:55 GMT
View Forum Message <> Reply to Message

I hope you will find answers on your questions here:
http://wiki.openvz.org/Privvmpages#privvmpages

thank you,
Vasily Averin

Subject: Re: Memory Management again
Posted by ledude1 on Thu, 03 Jan 2008 16:00:09 GMT
View Forum Message <> Reply to Message

Thanks Vasily.

It kind of explains it except that it didn't go into details on how it comes up with the 262140. Is there like a best practice number on how to obtain the number.

PRIVVMPAGES="256000:262140" // 1024MB Burst

Many thanks for taking your guys time to explain this number.

Subject: Re: Memory Management again
Posted by ledude1 on Mon, 07 Jan 2008 04:27:03 GMT
View Forum Message <> Reply to Message

#bump...

Anyone?

Subject: Re: Memory Management again
Posted by ugob on Mon, 07 Jan 2008 12:18:14 GMT

View Forum Message <> Reply to Message

It looks like they use a 1000:1024 ratio. This is probably just an arbitrary way to get human-readable parameters and give 'a little more' to the hard limit.

Subject: Re: Memory Management again
Posted by ledude1 on Mon, 07 Jan 2008 20:03:15 GMT
View Forum Message <> Reply to Message

Thanks Ugob.

I guess there's really no special number calculation that needs to be perform to get the number other than using the ratio you mentioned earlier. Very interesting. Thanks again Ugob.

Anyone else want to chime in?

Subject: Re: Memory Management again
Posted by vaverin on Tue, 08 Jan 2008 06:49:35 GMT
View Forum Message <> Reply to Message

http://wiki.openvz.org/Privvmpages#privvmpages:
"There should be a safety gap between the barrier and the limit for privvmpages parameter to reduce the number of memory allocation failures that the application is unable to handle. This gap will be used for “high-priority” memory allocations, such as process stack expansion.
Normal priority allocations will fail when the barrier if privvmpages is reached."

Subject: Re: Memory Management again
Posted by ledude1 on Tue, 08 Jan 2008 18:40:11 GMT
View Forum Message <> Reply to Message

Thanks Vasily.

What I'm missing is that how it ended up using the ratio 1000:1024 to obtain this number? Like I said earlier, i've read the manual but nowhere in it that stated a mathematical formula on how to get this number. 256000:262140.

But thanks again anyway.

Subject: Re: Memory Management again
Posted by vaverin on Wed, 09 Jan 2008 08:42:23 GMT
View Forum Message <> Reply to Message

In vzsplit utility we use 10% gap for privvmpages limit. However there is not any strong recommedation about this delta and you can use some % or just have some constant value for it.

thank you,
Vasily Averin

Subject: Re: Memory Management again
Posted by ledude1 on Wed, 09 Jan 2008 19:09:11 GMT
View Forum Message <> Reply to Message

Vasily.

Thanks. Now that answer my question. Its a guesstimate. and I'm fine with that. Many thanks again Vasily and keep up the good work everyone.
