
Subject: Calculating kmemsize
Posted by [spectator](#) on Sat, 18 Jul 2009 14:49:36 GMT
[View Forum Message](#) <> [Reply to Message](#)

I've been looking through the forums, wiki documentation and the examples on the forum but I've not been able to work out how to calculate kmemsize.

I'm beginning to think my understanding of how to calculate kmemsize is wrong.

From http://wiki.openvz.org/UBC_configuration_examples_table, I cannot work out how to get kmemsize.

Can someone explain how to do it?

Subject: Re: Calculating kmemsize
Posted by [spectator](#) on Sat, 18 Jul 2009 23:27:55 GMT
[View Forum Message](#) <> [Reply to Message](#)

I've been going through the source code for vzsplit:

<http://git.openvz.org/?p=vzctl;a=blob;f=src/vzsplit.c;h=33c9d1f54c4fa7b12438fa4aea94c42f540bfe8a;hb=68e1b402f09ae45b49df8d1ae22fd3bfb60324d4>

Unfortunately it's not been much help.

How do other people manage to create a container manually without knowing how to calculate kmemsize.

Subject: Re: Calculating kmemsize
Posted by [maratrus](#) on Mon, 20 Jul 2009 06:42:03 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Dave,

what does it mean "to calculate kmemsize"?

Do you want to calculate kmemsize before VE is created?

If you need to construct a function which calculates kmemsize you have to define what arguments your function takes.

Some consistency checks might be found here:

http://wiki.openvz.org/UBC_consistency_check

Don't you want to create an experimental kmemsize value i.e. at first make kmemsize (and other user_beancounters parameters) be unlimited then run inside VE whatever you want and then look at the usage of kmemsize (and other user_beancounters parameters) and adjust all properly.

Subject: Re: Calculating kmemsize

Posted by [spectator](#) on Mon, 20 Jul 2009 12:19:52 GMT

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

> what does it mean "to calculate kmemsize"?

I'm trying to work out how to obtain the value of kmemsize in the examples of the OpenVZ wiki.

> Do you want to calculate kmemsize before VE is created?

Yes, I do. For two reasons. 1) To ensure my understanding of the examples is correct and also so that I can create containers of varying sizes on the same node.

It looks like my understanding of kmemsize is wrong then.

Looking at the consistency check page on the OpenVZ wiki, I see the first equation states that kmemsize needs to be at least a certain size but I can't come to a value of 16,384,000 for kmemsize.

Here are the values I'm using:

```
total mem 248MB
kernel mem 42304KB
user mem 200MB
avnumproc 200
numproc 400
numtcpsock 500
numothersock 500
vmguarpages 102400
kmemsize 16,384,000 / 18,022,400
tcpsndbuf 5,365,760 / 10,485,760
tcprcvbuf 5,365,760 / 10,485,760
othersockbuf 1,503,232 / 4,063,232
dgramrcvbuf 262144
oomguarpages 102400
lockedpages 4096
shmpages 131072
privvmpages 262,144 / 292,912
numfile 8192
numflock 200 / 220
numpty 64
numsiginfo 512
```

dcachesize 4,194,304 / 4,317,184

When using the equation to calculate the minimum for kmemsize, I always get a smaller value than the examples.

Are the examples wrong?

Subject: Re: Calculating kmemsize
Posted by [maratrus](#) on Mon, 20 Jul 2009 12:35:59 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi,

Quote:

When using the equation to calculate the minimum for kmemsize, I always get a smaller value than the examples.

I cannot see any problem.

16,384,00 >= \$YOUR_KMEMSIZE >= \$FORMULA_VALUE

where is a contradiction?

Subject: Re: Calculating kmemsize
Posted by [spectator](#) on Mon, 20 Jul 2009 12:39:33 GMT
[View Forum Message](#) <> [Reply to Message](#)

There is no contradiction - it's more likely my understanding is wrong.

> 16,384,00 >= \$YOUR_KMEMSIZE >= \$FORMULA_VALUE

I currently get the value, 16,384,000 from the example page but have no idea how this number comes about. How do I get \$YOUR_KMEMSIZE or \$FORMULA_VALUE (I realise these are example variables)?

Or is it perhaps that I need to create a container then that GIVES me the value of kmemsize so kmemsize isn't calculated at all but is dependent on all the other variables of the system?

If so which variables combine to give kmemsize?

Subject: Re: Calculating kmemsize
Posted by [spectator](#) on Mon, 20 Jul 2009 12:57:56 GMT

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

For example, if I use figures from the OpenVZ example:

```
avnumproc = 200
datacachesize = 4,194,304
```

Using the consistency check formula:

```
kmemsize >= 40960 * avnumproc + datacachesize
```

Gives us: 12,386,304 not 16,384,000, as listed in the example.

I'm short 3,997,696 bytes.

Subject: Re: Calculating kmemsize
Posted by [seany](#) on Wed, 22 Jul 2009 14:31:33 GMT
[View Forum Message](#) <> [Reply to Message](#)

I too would like to know how to do this.

Right now we are "guessing" kmemsize - while we do satisfy the constraints we are wasting memory with wrong/generic kmemsize settings.

Subject: Re: Calculating kmemsize
Posted by [maratrus](#) on Wed, 22 Jul 2009 16:11:13 GMT
[View Forum Message](#) <> [Reply to Message](#)

There is no a precise formula.

It depends on applications that you're going to run inside the containers.
There is an lower bound which is mentioned on a wiki page so the upper bound may be obtained with help of natural experiment.

Subject: Re: Calculating kmemsize
Posted by [spectator](#) on Thu, 23 Jul 2009 10:14:21 GMT
[View Forum Message](#) <> [Reply to Message](#)

```
kmemsize >= 40960 * avnumproc + datacachesize
```

Would it be fair to say that this value is ok for 'average' use?

It strikes me as incredible that so few people need to know how to calculate this value.

Does everyone simply use vzsplit or is there some other method of provisioning a container?

I guess this would account for why there are so many badly configured containers out there.

Subject: Re: Calculating kmemsize

Posted by [damswf](#) on Wed, 02 Sep 2009 18:17:58 GMT

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

I agree, I just installed a web server in an openvz container, and i frequently have this problem. Either I can't do ps anymore, or i don't have a shell when I ssh, etc... instead I got "fork: Cannot allocate memory".

This is a pain for me since I wanted openvz to be more protected against hardware problems, and I have this headache instead.

Every time I grow up the kmemsize barrier limit, I see the maxheld increasing until being greater to the barrier limit then I got some failcnt and errors.

This server is not at all under load now, but once it will be, it will be because of a lot of customers hitting the site (I cannot reproduce the same thing myself to test), and I cannot take the risk to have databases crashing nor losing data because of being unable to allocate kernel memory on a 8GB 64bits system (vc have 4GB).

And I have no clue what the values should be! I have kmemsize $\geq 40kB * \text{avnumproc} + \text{dcache size}$ but still the problem.

This is a serious issue for production servers.

So what do I do? I explain to my boss I just virtualized our servers and we have more problems although I explained him a few times ago that it will work better?

Subject: Re: Calculating kmemsize

Posted by [dragos2](#) on Thu, 03 Sep 2009 13:34:01 GMT

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

damswf wrote on Wed, 02 September 2009 21:17: I agree, I just installed a web server in an openvz container, and i frequently have this problem. Either I can't do ps anymore, or i don't have a shell when I ssh, etc... instead I got "fork: Cannot allocate memory".

This is a pain for me since I wanted openvz to be more protected against hardware problems, and I have this headache instead.

Every time I grow up the kmemsize barrier limit, I see the maxheld increasing until being greater to the barrier limit then I got some failcnt and errors.

This server is not at all under load now, but once it will be, it will be because of a lot of customers hitting the site (I cannot reproduce the same thing myself to test), and I cannot take the risk to have databases crashing nor losing data because of being unable to allocate kernel memory on a 8GB 64bits system (vc have 4GB).

And I have no clue what the values should be! I have $\text{kmemsize} \geq 40\text{kB} * \text{avnumproc} + \text{dcache size}$ but still the problem.

This is a serious issue for production servers.

So what do I do? I explain to my boss I just virtualized our servers and we have more problems although I explained him a few times ago that it will work better?

You can try other virtualization types like Citrix FreeXenServer or even the opensource Xen(I don't have experience using Vmware ESX).

But using OpenVZ as a webserver in full load I think the main things to consider are:

- 1) number of processes and files opened
openvz parameter for processes number 66535 is a huge number
also the number of open file descriptors should be at least
twice the threads your webserver spawns
- 2) tcp/ip buffers size, number of sockets. Pay attention to these both on HN and as new container parameters. Be sure
to reuse them fast
- 3) cpu power if you make use of cgi, fcgi etc.

The upper values should be configured both on HN and in containers too.

If you set them only on one side, like only in HN or only in container you won't gain anything.

This was from my little experience.

I hope it helps.
