
Subject: fake swap != 0 in VE?

Posted by [Thomas Sattler](#) on Mon, 11 Feb 2008 18:23:16 GMT

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

Hi there ...

I'm running a software here, inside a VE,
that checks the amount of available swap
and complains if it thinks there isn't
enough.

The VE has plenty of RAM, swap will *never*
be used. But there are lots of popups about
missing swap all the time. :-(

Running 'free' in the VE shows no swap. Can
I fake some swapspace to prevent the popups?

Thomas

Subject: Re: fake swap != 0 in VE?

Posted by [Alex Prinsier](#) on Mon, 11 Feb 2008 18:26:55 GMT

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

1) I's say that software sucks. You better somehow disable the check,
even if it requires editing the code.

2) Perhaps you could use a file as swap instead of a normal swap
partition (google for that). You can make the size of that file really
small. However I'm really not so sure that you can mount a swap file in
openvz. It's worth a try.

Alexander

Thomas Sattler wrote:

> Hi there ...

>

> I'm running a software here, inside a VE,
> that checks the amount of available swap
> and complains if it thinks there isn't
> enough.

>

> The VE has plenty of RAM, swap will *never*
> be used. But there are lots of popups about
> missing swap all the time. :-(

>

> Running 'free' in the VE shows no swap. Can

> I fake some swapspace to prevent the popups?
>
> Thomas
>

Subject: Re: fake swap != 0 in VE?

Posted by [Thomas Sattler](#) on Mon, 11 Feb 2008 20:42:38 GMT

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

Alexander Prinsier wrote:

> 1) I's say that software sucks. You better somehow disable the check,
> even if it requires editing the code.

It's a closed source application, I can't do anything here. This really sucks. Perhaps its developers will have a look, perhaps they won't. :-/

> 2) Perhaps you could use a file as swap instead of a normal swap
> partition (google for that). You can make the size of that file really
> small. However I'm really not so sure that you can mount a swap file in
> openvz. It's worth a try.

I know about swapfiles. But swap is a system global thing and must be done in the HN not in the VE's:

```
# dd if=/dev/zero of=swap bs=1M count=32
32+0 records in
32+0 records out
33554432 bytes (34 MB) copied, 0.075684 seconds, 443 MB/s
```

```
# mkswap swap
Setting up swapspace version 1, size = 33550 kB
```

```
# swapon swap
swapon: swap: Operation not permitted
```

BTW: The HN uses swap:

```
# cat /proc/swaps
Filename                Type      Size  Used
Priority
/dev/md1                 partition 3903672 0    -1
```

It's only that the VE's don't know about that. So I need to tell them -- or fake it. That's what I asked for.

Thomas

Subject: Re: fake swap != 0 in VE?

Posted by [Michael Klatsky](#) on Mon, 11 Feb 2008 21:08:12 GMT

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

We have, unfortunately, run into a similar situation.

Our app requires swap as well, and, while you can set up a VE such that issuing 'free' will show both RAM and swap- it will be that of the HN. And, unfortunately, our app requires swap to be seen as part of the VE, as in issuing a 'df'. I have tried making a swapfile and mounting (swapon) with the same errors. It doesn't appear possible for the VE to really show swap. Unfortunately, for our particular app, we may need to use XEN.

On 2/11/08, Thomas Sattler <sattler@med.uni-frankfurt.de> wrote:

> Alexander Prinsier wrote:

> > 1) I's say that software sucks. You better somehow disable the check,
> > even if it requires editing the code.

>

> It's a closed source application, I can't do anything here. This
> really sucks. Perhaps its developers will have a look, perhaps
> they won't. :-/

>

> > 2) Perhaps you could use a file as swap instead of a normal swap
> > partition (google for that). You can make the size of that file really
> > small. However I'm really not so sure that you can mount a swap file in
> > openvz. It's worth a try.

>

> I know about swapfiles. But swap is a system global thing and
> must be done in the HN not in the VE's:

>

> # dd if=/dev/zero of=swap bs=1M count=32

> 32+0 records in

> 32+0 records out

> 33554432 bytes (34 MB) copied, 0.075684 seconds, 443 MB/s

>

> # mkswap swap

> Setting up swspace version 1, size = 33550 kB

>

> # swapon swap

> swapon: swap: Operation not permitted

>

> BTW: The HN uses swap:

>

> # cat /proc/swaps

Filename	Type	Size	Used
----------	------	------	------

Priority			
----------	--	--	--

/dev/md1	partition	3903672 0	-1
----------	-----------	-----------	----

>

> It's only that the VE's don't know about that. So I need to
> tell them -- or fake it. That's what I asked for.
>
> Thomas
--

Sincerely,

Michael Klatsky

Subject: Re: fake swap != 0 in VE?

Posted by [Thomas Sattler](#) on Tue, 12 Feb 2008 08:03:23 GMT

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

> [...] while you can set up a VE such that issuing 'free'
> will show both RAM and swap- it will be that of the HN.

That might be all that I need. I can't tell how
the application checks for swap. Could someone
give me a hint how an application (like 'free')
can see any swap?

I'm quite sure that I do not *need* the swap as
there is plenty of RAM. But the application com-
plains if there is no swap to be seen.

Thomas

--

Dipl.-Inform.
Thomas Sattler
Sysadmin/DVK
BIC Kernstruktur

Uniklinikum FFM
ZNN/BIC (Haus 95H)
Schleusenweg 2-16
D-60528 Frankfurt

Tel: +49 69 6301 6895
Fax: +49 69 6301 3707

Subject: Re: fake swap != 0 in VE?

Posted by [Kirill Korotaev](#) on Tue, 12 Feb 2008 08:46:28 GMT

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

the only way I'm aware of is /proc/meminfo
just patch kernel to print something non-zero there always and your application most likely will become happy.

Thomas Sattler wrote:

>> [...] while you can set up a VE such that issuing 'free'
>> will show both RAM and swap- it will be that of the HN.

>

> That might be all that I need. I can't tell how
> the application checks for swap. Could someone
> give me a hint how an application (like 'free')
> can see any swap?

>

> I'm quite sure that I do not *need* the swap as
> there is plenty of RAM. But the application com-
> plains if there is no swap to be seen.

>

> Thomas

>

Subject: Re: fake swap != 0 in VE?

Posted by [Thomas Sattler](#) on Tue, 12 Feb 2008 09:31:46 GMT

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

Kirill Korotaev wrote:

> the only way I'm aware of is /proc/meminfo
> just patch kernel to print something non-zero
> there always and your application most likely
> will become happy.

So far, I used kernels via

deb <http://download.openvz.org/debian-systs> etch openvz

Is there a chance that this becomes a tunable in the
"official openvz-kernels"?

Thomas

Subject: Re: fake swap != 0 in VE?

Posted by [Kirill Korotaev](#) on Tue, 12 Feb 2008 09:43:44 GMT

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

if someone cares to create patch for kernel/vzctl, why not?

Thomas Sattler wrote:

> Kirill Korotaev wrote:

>> the only way I'm aware of is /proc/meminfo

>> just patch kernel to print something non-zero

>> there always and your application most likely

>> will become happy.

>

> So far, I used kernels via

>

> deb <http://download.openvz.org/debian-systs> etch openvz

>

> Is there a chance that this becomes a tunable in the

> "official openvz-kernels"?

>

> Thomas

Subject: Re: fake swap != 0 in VE?

Posted by [Carl-Daniel Hailfinger](#) on Tue, 12 Feb 2008 12:13:02 GMT

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

On 12.02.2008 09:46, Kirill Korotaev wrote:

> the only way I'm aware of is /proc/meminfo

> just patch kernel to print something non-zero there always and your application most likely will become happy.

>

Why not

mount --bind fakememinfo /proc/meminfo

inside the VE? Would be easiest, but the contents will not match real meminfo because they do not get updated.

Regards,

Carl-Daniel

--

<http://www.hailfinger.org/>

Subject: Re: fake swap != 0 in VE? (solved)

Posted by [Thomas Sattler](#) on Wed, 13 Feb 2008 09:50:49 GMT

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

> Why not

> mount --bind fakememinfo /proc/meminfo
> inside the VE? Would be easiest, but the
> contents will not match real meminfo
> because they do not get updated.

I didn't expect this to work ...
... But it does. :-)

I attached a little script to be run as root
inside a VE to fake swap. 512MB are default,
you can change it on the commandline:

Output is, in the example above, cut at "free"
to prevent linewrap:

```
# free -m
```

	total	used	free
Mem:	4096	104	3991
-/+ buffers/cache:		104	3991
Swap:	0	0	0

```
# openvz_fake_swap
```

	total	used	free
Mem:	4096	104	3991
-/+ buffers/cache:		104	3991
Swap:	512	0	512

```
# openvz_fake_swap 4096
```

	total	used	free
Mem:	4096	104	3991
-/+ buffers/cache:		104	3991
Swap:	4096	0	4096

Thanks a lot!
Thomas

```
#!/bin/bash
```

```
SWAP="${1:-512}"
```

```
NEW="$[SWAP*1024]"; TEMP="${NEW//?/ }"; OLD="${TEMP:1}0"
```

```
umount /proc/meminfo 2> /dev/null  
sed "/^Swap(Total|Free):/s,$OLD,$NEW," /proc/meminfo > /etc/fake_meminfo  
mount --bind /etc/fake_meminfo /proc/meminfo
```

```
free -m
```

Subject: Re: fake swap != 0 in VE? (solved)
Posted by [Dariush Pietrzak](#) on Wed, 13 Feb 2008 13:13:15 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Wed, 13 Feb 2008, Thomas Sattler wrote:

> >Why not
> >mount --bind fakememinfo /proc/meminfo
> >inside the VE? Would be easiest, but the
> >contents will not match real meminfo
> >because they do not get updated.
>
> I didn't expect this to work ...
> ... But it does. :-)
>
> I attached a little script to be run as root
> inside a VE to fake swap. 512MB are default,
> you can change it on the commandline:

Hmm, I find it strange that such hacks are needed, especially when openvz already does do some kind of virtualization of /proc/meminfo (if I'm interpreting this correctly, it sets max-mem from userbeancounters barrier on VE start, but subsequent online changes are ignored)

--

Key fingerprint = 40D0 9FFB 9939 7320 8294 05E0 BCC7 02C4 75CC 50D9
Total Existence Failure

Subject: Re: fake swap != 0 in VE? (solved)
Posted by [Michael Klatsky](#) on Wed, 13 Feb 2008 14:24:31 GMT
[View Forum Message](#) <> [Reply to Message](#)

Would not the same effect be achived using OpenVZ builtin --meminfo parameter to vzctl?

Example:

Before:

```
[root@vz-210 /]# free -m
```

	total	used	free	shared	buffers	cached
Mem:	4050	21	4028	0	0	0
-/+ buffers/cache:		21	4028			
Swap:	0	0	0			

Then, do:

vzctl set 210 --meminfo=none

Now:

```
[root@vz-210 /]# free -m
```

	total	used	free	shared	buffers	cached
Mem:	4050	1963	2087	0	84	1655
-/+ buffers/cache:		223	3826			
Swap:	2047	0	2047			

Good discussion as I am having issues with certain applications that apparently need swap.

MK

On 2/13/08, Dariush Pietrzak <ml-openvz-eyck@ddiary.eu> wrote:

> On Wed, 13 Feb 2008, Thomas Sattler wrote:

>

> > >Why not

> > >mount --bind fakememinfo /proc/meminfo

> > >inside the VE? Would be easiest, but the

> > >contents will not match real meminfo

> > >because they do not get updated.

> >

> > I didn't expect this to work ...

> > ... But it does. :-)

> >

> > I attached a little script to be run as root

> > inside a VE to fake swap. 512MB are default,

> > you can change it on the commandline:

>

> Hmm, I find it strange that such hacks are needed, especially when openvz

> already does do some kind of virtualization of /proc/meminfo (if I'm

> interpreting this correctly, it sets max-mem from userbeancounters barrier

> on VE start, but subsequent online changes are ignored)

> --

> Key fingerprint = 40D0 9FFB 9939 7320 8294 05E0 BCC7 02C4 75CC 50D9

> Total Existence Failure

--

Sincerely,

Michael Klatsky

Subject: Re: fake swap != 0 in VE? (solved)

Posted by [Dariush Pietrzak](#) on Wed, 13 Feb 2008 14:52:35 GMT

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

> Would not the same effect be achieved using OpenVZ builtin --meminfo

> parameter to vzctl?

oooh, this is good,

however I can't find any info/docs about this option, the only thing I found is:

"Recently Vasily Averin very clearly explained the reason but it was in Russian part of OpenVZ forum. Smile

I'll try to reproduce his idea very precisely.

The fact is that you can restrict the userspace memory via privmpages but inside VE we cannot divide it among the SWAP and RAM. In theory all VE's userspace memory can be swapped that is why we have to represent the swap size inside VE like privmpages.

But also VE's userspace can stay in RAM so we have to take into account privmpages in RAM accounting. So we have take into account privvmpages in RAM and in SWAP accounting. That is why inside VE swap is shown like zero. But if your application uses the top/free output you can switch off /proc/meminfo virtualization by using vzctl's --meminfo parameter. And get the same /proc/meminfo like on host system."

--

Key fingerprint = 40D0 9FFB 9939 7320 8294 05E0 BCC7 02C4 75CC 50D9

Total Existence Failure
