
Subject: Re: Limits for low memory

Posted by [dev](#) on Wed, 25 Apr 2007 14:19:38 GMT

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

Wolfgang Schnerring wrote:

> Hello!

>

> I've been doing some experiments with OpenVZ, and one of them was, how
> many VE's can I run on a single machine? The guinea pig machine was a
> Pentium4/3GHz with 2GB RAM. 230 VE's run just fine, but upon starting
> the 231st, I reproducably get a kernel panic, "out of memory" (VE
> configuration is appended below).

1. 230 VEs correspond to kir@ results when he was running fine ~120 VEs on 1Gb notebook.
2. can you please attach your kernel panic
(if it was a really panic and not an OOM kill message)?
3. what kernel version do you use?

> What I'd really like to know is, why does that happen?

panic or ?

> So I collected resource consumption numbers from /proc/bc for all VE's.
> Summed up for 220 VE's (I didn't dare 230 ;-), the maxheld values are
> kmemsize 160957978
> socket buffers 167838032 (tcp_sndbuf+tcp_rcvbuf+dgram_rcvbuf+othersockbuf)
> physpages 143708
>
> The physpages amount to 562MB; that should fit into 2GB nicely in my book.

is it i686 or x86-64 kernel?

actually it is not all.

1. You also need to take into account dcachesize
(which is not immediately accounted by default for optimization purposes,
but can be enabled with dentry_watermark sysctl and
usually dcachesize is around ~2Mb per VE, so 460Mb of low memory in your case)
2. plus there are some in-kernel structures which are not-accounted in kmemsize,
since can't be used for DoS. i.e. kmemsize accounts only those resources
which can be used for abusing. But real usage of kernel memory can be slightly higher.

> If I understand <http://wiki.openvz.org/UBC_systemwide_configuration>
> correctly, those two things need to go into low memory, and the
> total amount for that on x86 is 832MB.

No. lowmemory is required for: kmemsize, tcp/socket/udp buffers and dcachesize

physpages can go both to low memory and highmemory and you don't know how it was spread from UBC counters. (*)

> -- Question: Where do the 832 MB come from?

32bit systems can address up to 4GB of virtual RAM (2^{32} bytes).
this 4GB address space is split by default as 3 GB for user space
and 1GB for kernel space.

1GB of kernel has kernel itself, different reserved areas (e.g. for vmalloc by default 128Mb),
So around ~850Mb of address space is left for free memory.

The number depends on kernel and configuration:

```
[dev@dev ~]$ cat /proc/meminfo | grep Low  
LowTotal:      903348 kB
```

the most impact on the size of low memory have vmalloc reserved area.

We used to reserve more when different kernel structures were using it, so low memory was smaller.

Nowdays only 128Mb are reserved by default and about ~890Mb of Low memory should be available.

> The two numbers above are roughly 314 MB, that's nowhere near 832.

> But the wiki-page has the formula $\text{kmemsize} + \text{sockbuf} / 0.4 \cdot 832\text{MB}$ should

> be smaller than 1 to be safe. Interestingly enough, that quotient is
> 0.94, so there might be something up here.

>

> -- Question: Where does the 0.4 come from?

in your case:

kmemsize ~160Mb

buffers ~160Mb

dcachesize ~460Mb

i.e. ~780Mb in total (see also (*)).

plus you need to take into account fragmentation issues.

e.g. if some memory page was allocated for objects of size 128 bytes,
and has only one such object allocated on it, the page won't be used for objects of another size
and thus some of the memory will be unused and can't be used even under memory pressure.

i.e. it's not that easy like a sum of counters == total memory.

> Can anybody help me figure out what's going on here?

you are welcome!

> Thank you very much for your help,

```
> Wolfgang
>
> # VE configuration file
> NUMFLOCK="100:110"
> VMGUARPAGES="6144:2147483647"
> OTHERSOCKBUF="132096:459776"
> DISTRIBUTION="debian"
> NUMPTY="16:16"
> OOMGUARPAGES="6144:2147483647"
> DGRAMRCVBUF="132096:132096"
> CPUUNITS="1000"
> NUMIPTENT="128:128"
> NUMTCPSOCK="768:768"
> PHYSPAGES="0:2147483647"
> AVNUMPROC="50:50"
> NUMPROC="100:100"
> LOCKEDPAGES="32:32"
> TCPSNDBUF="10485760:12451840"
> VE_ROOT="/var/tmp/vte-wosc/vm33745/root"
> NUMSIGINFO="256:256"
> KMEMSIZE="5324800:5857280"
> SHMPAGES="8192:8192"
> TCPRCVBUF="10485760:12451840"
> VE_PRIVATE="/var/tmp/vte-wosc/vm33745/private"
> NUMFILE="3200:3200"
> PRIVVMPAGES="32768:36044"
> DCACHESIZE="1228800:1351680"
> NUMOTHERSOCK="128:128"
> OSTEMPLATE="debian-3.1"
> HOSTNAME="vm0.local"
> NETIF=" ifname=eth0,mac=02:7B:6A:56:00:02,host_ifname=vif33745.0,host_
t_mac=02:7B:6A:56:00:02 "
Kirill
```
