
Subject: ovzkernel-xen 4gb fix up issue

Posted by [victorski](#) on Fri, 16 Jan 2009 04:34:31 GMT

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

Hi,

I found when running ovzkernel-xen based kernel, 4gb fix up issue could not solve out. Here are a few report.

```
[root@firefly ~]# yum list installed | grep kernel
kernel-headers.i386          2.6.18-92.1.22.el5    installed
kernel-xen.i686             2.6.18-92.1.22.el5    installed
kernel-xen.i686             2.6.18-92.1.13.el5    installed
kernel-xen.i686             2.6.18-92.el5         installed
kernel-xen-devel.i686       2.6.18-92.1.22.el5    installed
ovzkernel-xen.i686          2.6.18-92.1.13.el5.028 installed
```

```
[root@firefly ~]# yum list installed | grep glibc
glibc.i686                  2.5-24.el5_2.2        installed
glibc-common.i386           2.5-24.el5_2.2        installed
glibc-devel.i386            2.5-24.el5_2.2        installed
glibc-headers.i386          2.5-24.el5_2.2        installed
```

Note that, each kernel-xen created `hwcap 0 nosegnet` bit, since all are rhel base kernels and red hat include this xen bit patch.

```
[root@firefly ~]# ls /etc/ld.so.conf.d/
kernelcap-2.6.18-92.1.13.el5.028stab059.6.conf kernelcap-2.6.18-92.el5.conf
kernelcap-2.6.18-92.1.13.el5.conf  mysql-i386.conf
kernelcap-2.6.18-92.1.22.el5.conf
[root@firefly ~]# cat /etc/ld.so.conf.d/kernelcap-*
# This directive teaches ldconfig to search in nosegnet subdirectories
# and cache the DSOs there with extra bit 0 set in their hwcap match
# fields. In Xen guest kernels, the vDSO tells the dynamic linker to
# search in nosegnet subdirectories and to match this extra hwcap bit
# in the ld.so.cache file.
hwcap 0 nosegnet
# This directive teaches ldconfig to search in nosegnet subdirectories
# and cache the DSOs there with extra bit 0 set in their hwcap match
# fields. In Xen guest kernels, the vDSO tells the dynamic linker to
# search in nosegnet subdirectories and to match this extra hwcap bit
# in the ld.so.cache file.
hwcap 0 nosegnet
# This directive teaches ldconfig to search in nosegnet subdirectories
```

```
# and cache the DSOs there with extra bit 0 set in their hwcap match
# fields. In Xen guest kernels, the vDSO tells the dynamic linker to
# search in nosegneg subdirectories and to match this extra hwcap bit
# in the ld.so.cache file.
```

```
hwcap 0 nosegneg
```

```
# This directive teaches ldconfig to search in nosegneg subdirectories
# and cache the DSOs there with extra bit 0 set in their hwcap match
# fields. In Xen guest kernels, the vDSO tells the dynamic linker to
# search in nosegneg subdirectories and to match this extra hwcap bit
# in the ld.so.cache file.
```

```
hwcap 0 nosegneg
```

When running with RHEL5.2 latest 22 kernel, it run fine. No error printk. libc.so.6 point to nosegneg.

```
[root@firefly ~]$ uname -r
2.6.18-92.1.22.el5xen
[root@firefly ~]$ ldd /sbin/init
linux-gate.so.1 => (0x00a01000)
libsepol.so.1 => /lib/libsepol.so.1 (0x00b99000)
libseline.so.1 => /lib/libseline.so.1 (0x00b7f000)
libc.so.6 => /lib/i686/nosegneg/libc.so.6 (0x00110000)
libdl.so.2 => /lib/libdl.so.2 (0x00b37000)
/lib/ld-linux.so.2 (0x009d1000)
```

When running with RHEL5 13 kernel(where ovzkernel-xen 2.6.18-92.1.13.el5.028stab059.6xen base off), it also run fine. No error printk. libc.so.6 point to nosegneg.

```
[root@firefly ~]# uname -r
2.6.18-92.1.13.el5xen
[root@firefly ~]# ldd /sbin/init
linux-gate.so.1 => (0x00265000)
libsepol.so.1 => /lib/libsepol.so.1 (0x00b99000)
libseline.so.1 => /lib/libseline.so.1 (0x00b7f000)
libc.so.6 => /lib/i686/nosegneg/libc.so.6 (0x009ef000)
libdl.so.2 => /lib/libdl.so.2 (0x00b37000)
/lib/ld-linux.so.2 (0x009d1000)
```

Once I installed ovzkernel-xen and boot on, it fill the console with 4gb fix up printk messages. libc.so.6 _do not_ point to nosegneg, anymore.

```
[root@firefly ~]# uname -r
2.6.18-92.1.13.el5.028stab059.6xen
[root@firefly ~]# ldd /sbin/init
```

libsepol.so.1 => /lib/libsepol.so.1 (0x00b99000)
libselinux.so.1 => /lib/libselinux.so.1 (0x00b7f000)
libc.so.6 => /lib/libc.so.6 (0x00a1e000)
libdl.so.2 => /lib/libdl.so.2 (0x00110000)
/lib/ld-linux.so.2 (0x009d1000)

As soon as node boot up and start init pid 1, then it started 4gb fix up printk.

Quote:Console Messages

.....

.....

Setting up other filesystems.

Setting up new root fs

no fstab.sys, mounting internal defaults

Switching to new root and running init.

unmounting old /dev

unmounting old /proc

unmounting old /sys

4gb seg fixup, process init (pid 1), cs:ip 73:004ae5a0

4gb seg fixup, process init (pid 1), cs:ip 73:00524939

4gb seg fixup, process init (pid 1), cs:ip 73:004b16cc

4gb seg fixup, process init (pid 1), cs:ip 73:004b16da

4gb seg fixup, process init (pid 1), cs:ip 73:004b16cc

4gb seg fixup, process init (pid 1), cs:ip 73:004b16da

4gb seg fixup, process init (pid 1), cs:ip 73:00469a8b

4gb seg fixup, process init (pid 1), cs:ip 73:00469a74

4gb seg fixup, process init (pid 1), cs:ip 73:004b77e5

4gb seg fixup, process init (pid 1), cs:ip 73:00469a74

INIT: version 2.86 booting

Welcome to CentOS release 5.2 (Final)

Press 'l' to enter interactive startup.

printk: 115503 messages suppressed.

4gb seg fixup, process hwclock (pid 766), cs:ip 73:0014b810

.....

.....

<snip>

.....

4gb seg fixup, process sendmail (pid 4343), cs:ip 73:00ce06cc

printk: 16 messages suppressed.

4gb seg fixup, process sendmail (pid 4343), cs:ip 73:00ce06cc

printk: 18 messages suppressed.

4gb seg fixup, process sendmail (pid 4343), cs:ip 73:00ce06cc

printk: 16 messages suppressed.

4gb seg fixup, process sendmail (pid 4343), cs:ip 73:00ce06cc

printk: 16 messages suppressed.

4gb seg fixup, process sendmail (pid 4343), cs:ip 73:00ce06cc

printk: 15 messages suppressed.

The error is reproducible. I also found others' happening on this list thread.
<http://openvz.org/pipermail/users/2008-December/002462.html>

I've tried to rebuild from <http://wiki.openvz.org/Download/kernel/rhel5/028stab059.6> source rpm but still not work. I found the following threads might give some help. It would be related to OpenVZ code to align with how Glibc treat on Xen's TLS/ELF segment, regarding syscall to user space linker `_but_` not so sure if my understanding is correct. I've tried to check in ovzkernel code but still new in kernel hacking and openvz code is a big patch also, diff about 800++ files on rhel base.

http://lists.xensource.com/archives/html/xen-devel/2005-11/m_sg00262.html
http://lists.xensource.com/archives/html/xen-devel/2005-11/m_sg00078.html
<http://lkml.indiana.edu/hypermail/linux/kernel/0704.3/0002.html>
<https://lists.linux-foundation.org/pipermail/virtualization/2007-May/007631.html>

Please kindly help prior on this issue, as message logs is growing up and kind of effecting in performance, if keep stubborn to run with this kernel.

```
[root@node ~]# ls -alh /var/log/messages*  
-rw----- 1 root root 15M Jan 16 12:39 /var/log/messages  
-rw----- 1 root root 19M Jan 11 04:02 /var/log/messages.1  
-rw----- 1 root root 9.2M Jan 4 04:02 /var/log/messages.2  
-rw----- 1 root root 34K Dec 28 04:02 /var/log/messages.3  
-rw----- 1 root root 294K Dec 21 04:02 /var/log/messages.4
```

Thanks.

P.S.. This regardless of Dom0(Host) or DomU(Guest), which ever running this kernel will get printk!
