

---

Subject: \*SOLVED\* /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [John Kelly](#) on Fri, 10 Nov 2006 21:27:43 GMT

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

---

I have this error when restarting sendmail in a suse 9.1 VE:

startproc: cannot stat /proc/27759/exe

I see in /proc the sendmail pid number is one higher, 27760.

Is this an off-by-one 2.6.18 bug? I never had this problem while running a 2.6.16 openvz kernel.

---

Subject: Re: /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [Vasily Tarasov](#) on Sun, 12 Nov 2006 14:25:11 GMT

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

---

Hello,

can you somehow figure out where it gets 27759 pid?

I guess, strace can be used to find this information.

One more question is about other templates? Does this problem present in other than Suse templates?

Thanks!

---

Subject: Re: /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [John Kelly](#) on Mon, 13 Nov 2006 17:19:59 GMT

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

---

#!/bin/sh

SENDMAIL\_CLIENT\_ARGS="-L sendmail-client -Ac -qp30m"

msppid=/var/spool/clientmqueue/sm-client.pid

srvpid=/var/run/sendmail.pid

killproc -p \$msppid -i \$srvpid -TERM /usr/sbin/sendmail

startproc -p \$msppid -i \$srvpid /usr/sbin/sendmail \$SENDMAIL\_CLIENT\_ARGS

Here is a reduced test case, the problem happens on the last line, startproc. The problem seems like some kind of race, because sometimes it happens, and other times, it does not.

I tried strace with startproc, but that seems to avoid the race. However, after running the test script above many times, followed immediately by "ps ax," I was able to see what the problem is (shown below). There is a zombie with the PID number in question, and the actual PID number of the running sendmail process is one higher. Seeing the zombie with "ps ax" is hard to reproduce, I only captured it one time.

This never happened until I started using the openvz 2.6.18 kernel. I don't know if this happens with any other VE, suse 9.1 is the only one I use enough to produce the problem.

startproc: cannot stat /proc/1372/exe: Permission denied

| PID   | TTY   | STAT | TIME | COMMAND  |
|-------|-------|------|------|--|
| 1     | ?     | Rs   | 0:00 | init [3]   |
| 28095 | ?     | Ss   | 0:00 | sendmail: accepting connections                  |
| 28107 | ?     | Ss   | 0:00 | /usr/sbin/sshd -o PidFile=/var/run/sshd.init.pid |
| 28113 | ?     | Ss   | 0:00 | /usr/sbin/xinetd                                 |
| 28119 | ?     | Ss   | 0:00 | /usr/sbin/cron                                   |
| 28276 | pts/1 | Ss+  | 0:00 | -bash  |
| 1372  | pts/0 | Z    | 0:00 | [sendmail] <defunct>                             |
| 1373  | ?     | Ss   | 0:00 | sendmail: Queue control                          |
| 1374  | ?     | S    | 0:00 | sendmail: running queue: /var/spool/clientmqueue |
| 1375  | pts/0 | R+   | 0:00 | ps ax  |

---

Subject: Re: /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [dev](#) on Tue, 14 Nov 2006 16:50:20 GMT

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

---

1. can you run it as:

```
# bash -x <your-script>
please?
```

2. AFAICS, this script does the following:

a) kills previous sendmail instance

b) starts new sendmail instance

however, the problem is that SIGTERM requires some time to finish.

So from your example:

```
1372 pts/0  Z   0:00 [sendmail] <defunct>
is an old sendmail instance.
```

```
1373 ?     Ss   0:00 sendmail: Queue control
1374 ?     S   0:00 sendmail: running queue: /var/spool/clientmqueue
a new one.
```

and looks like startproc races with SIGTERM and sees that the task still exists, however, when it tries to do stat on /proc/pid/exe it is already dead and it can't stat.

Looks like this.

---

---

Subject: Re: /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [John Kelly](#) on Tue, 14 Nov 2006 17:35:23 GMT

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

---

A race between SIGTERM and startproc is what I thought too. But then I put "sleep 1" between killproc and startproc, and the problem still happens.

The zombie pid is not the old pid that was killed; the old pid number is much lower (not just one lower); I can see that before killing it. The zombie pid is somehow related to the new instance of sendmail, though I am not sure how.

As another data point:

This morning I was using a debian etch VE, running aptitude interactively to install a package. But after downloading, it stalled, and went no further. Then I used another session to look with "ps ax" and I saw another zombie:

| PID   | TTY   | STAT | TIME | COMMAND   |
|-------|-------|------|------|---|
| 1     | ?     | Ss   | 0:00 | init [2]  |
| 9359  | ?     | Ss   | 0:00 | /sbin/syslogd   |
| 9365  | ?     | Ss   | 0:00 | /sbin/klogd -x  |
| 9377  | ?     | Ssl  | 0:00 | /usr/sbin/named -u bind                               |
| 9401  | ?     | Ss   | 0:00 | /usr/sbin/sshd  |
| 9405  | ?     | Ss   | 0:00 | /usr/sbin/vsftpd                                      |
| 9411  | ?     | Ss   | 0:00 | /usr/sbin/xinetd -pidfile /var/run/xinetd.pid -stayal |
| 9432  | ?     | Ss   | 0:00 | /usr/sbin/cron  |
| 19964 | ?     | Ss   | 0:00 | sshd: root@pts/0                                      |
| 19967 | pts/0 | Ss   | 0:00 | -bash   |
| 21510 | ?     | Ss   | 0:00 | sshd: root@pts/2                                      |
| 21513 | pts/2 | Ss   | 0:00 | -bash   |
| 31876 | pts/2 | Zl+  | 0:08 | [aptitude] <defunct>                                  |
| 31964 | pts/0 | R+   | 0:00 | ps ax   |

This never happened to me before, with debian aptitude. Maybe there is some 2.6.18 kernel regression related to PIDs and zombies, but I don't know how to analyze it further.

---

Subject: Re: /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [dev](#) on Tue, 14 Nov 2006 22:50:40 GMT

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

---

can you give me an access to the node with exact instructions on reproducing both issues (sendmail and aptitude)?

---

---

Subject: Re: /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [John Kelly](#) on Wed, 15 Nov 2006 02:33:12 GMT

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

---

dev wrote on Tue, 14 November 2006 17:50 can you give me an access to the node with exact instructions on reproducing both issues (sendmail and aptitude)?

Yes for aptitude, since that's in a test environment. Please send me an email, and tell me your IP address. I protect ssh logins with /etc/hosts.allow.

Email: [jak@isp2dial.com](mailto:jak@isp2dial.com)

Alternate email: [isp2dial@fastmail.fm](mailto:isp2dial@fastmail.fm)

My kernel config is kernel-2.6.18-028test003-i686.config.ovz, with local changes, mostly to drop unneeded network and scsi drivers. I did remove the VDSO compat, but according to what I read, that should not make any difference, since my glibc is new enough.

Here are my kernel config changes which may possibly be relevant:

```
--- kernel-2.6.18-028test003-i686.config.ovz  2006-11-09 12:33:27.000000000 -0500
+++ k2618.openvz.v1      2006-11-10 09:23:23.000000000 -0500
@@ -1,7 +1,7 @@
#
# Automatically generated make config: don't edit
# Linux kernel version: 2.6.18-028test003
-# Thu Nov  9 17:34:51 2006
+# Fri Nov 10 09:23:23 2006
#
CONFIG_X86_32=y
CONFIG_GENERIC_TIME=y
@@ -194,14 +194,14 @@
# CONFIG_EFI is not set
# CONFIG_REGPARM is not set
# CONFIG_SECCOMP is not set
-# CONFIG_HZ_100 is not set
+CONFIG_HZ_100=y
# CONFIG_HZ_250 is not set
-CONFIG_HZ_1000=y
-CONFIG_HZ=1000
+# CONFIG_HZ_1000 is not set
+CONFIG_HZ=100
# CONFIG_KEXEC is not set
# CONFIG_CRASH_DUMP is not set
CONFIG_PHYSICAL_START=0x100000
-CONFIG_COMPAT_VDSO=y
+# CONFIG_COMPAT_VDSO is not set
CONFIG_ARCH_ENABLE_MEMORY_HOTPLUG=y

#
```

@ @ -858,7 +827,7 @ @

#

CONFIG\_NETDEVICES=y

CONFIG\_DUMMY=m

-CONFIG\_BONDING=m

+# CONFIG\_BONDING is not set

# CONFIG\_EQUALIZER is not set

CONFIG\_TUN=m

@ @ -1111,8 +1068,56 @ @

#

# Watchdog Cards

#

-# CONFIG\_WATCHDOG is not set

-# CONFIG\_HW\_RANDOM is not set

+CONFIG\_WATCHDOG=y

+# CONFIG\_WATCHDOG\_NOWAYOUT is not set

+

+#

+# Watchdog Device Drivers

+#

+CONFIG\_SOFT\_WATCHDOG=m

+# CONFIG\_ACQUIRE\_WDT is not set

+# CONFIG\_ADVANTECH\_WDT is not set

+# CONFIG\_ALIM1535\_WDT is not set

+# CONFIG\_ALIM7101\_WDT is not set

+# CONFIG\_SC520\_WDT is not set

+# CONFIG\_EUROTECH\_WDT is not set

+# CONFIG\_IB700\_WDT is not set

+# CONFIG\_IBMASR is not set

+# CONFIG\_WAFER\_WDT is not set

+# CONFIG\_I6300ESB\_WDT is not set

+CONFIG\_I8XX\_TCO=m

+# CONFIG\_SC1200\_WDT is not set

+# CONFIG\_60XX\_WDT is not set

+# CONFIG\_SBC8360\_WDT is not set

+# CONFIG\_CPU5\_WDT is not set

+# CONFIG\_W83627HF\_WDT is not set

+# CONFIG\_W83877F\_WDT is not set

+# CONFIG\_W83977F\_WDT is not set

+# CONFIG\_MACHZ\_WDT is not set

+# CONFIG\_SBC\_EPX\_C3\_WATCHDOG is not set

+

+#

+# ISA-based Watchdog Cards

+#

+# CONFIG\_PCWATCHDOG is not set

+# CONFIG\_MIXCOMWD is not set

```
+# CONFIG_WDT is not set
+
+#
+# PCI-based Watchdog Cards
+#
+# CONFIG_PCIPCWATCHDOG is not set
+# CONFIG_WDTPCI is not set
+
+#
+# USB-based Watchdog Cards
+#
+# CONFIG_USBPCWATCHDOG is not set
+CONFIG_HW_RANDOM=y
+CONFIG_HW_RANDOM_INTEL=m
+CONFIG_HW_RANDOM_AMD=m
+# CONFIG_HW_RANDOM_GEODE is not set
+CONFIG_HW_RANDOM_VIA=m
# CONFIG_NVRAM is not set
CONFIG_RTC=y
# CONFIG_DTLK is not set
@@ -1492,10 +1497,7 @@
CONFIG_JBD=y
CONFIG_JBD_DEBUG=y
CONFIG_FS_MBCACHE=y
-CONFIG_REISERFS_FS=y
-# CONFIG_REISERFS_CHECK is not set
-CONFIG_REISERFS_PROC_INFO=y
-# CONFIG_REISERFS_FS_XATTR is not set
+# CONFIG_REISERFS_FS is not set
# CONFIG_JFS_FS is not set
# CONFIG_FS_POSIX_ACL is not set
# CONFIG_XFS_FS is not set
```

---

Subject: Re: /proc pid number off-by-one? ... 2.6.18-028test003.1

Posted by [dev](#) on Mon, 20 Nov 2006 12:04:58 GMT

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

---

see the bug details for the patch:

[http://bugzilla.openvz.org/show\\_bug.cgi?id=352](http://bugzilla.openvz.org/show_bug.cgi?id=352)

---