

pfSense - Bug #7925

VT race condition panic at boot on ESXi 6.5.0U1 and FreeBSD 11.1 base

10/11/2017 09:21 AM - Jim Pingle

Status:	Resolved	Start date:	10/11/2017
Priority:	Normal	Due date:	
Assignee:	Luiz Souza	% Done:	100%
Category:	Operating System	Estimated time:	0.00 hour
Target version:	2.4.1	Affected Architecture:	amd64
Affected Version:	2.4.x		

Description

Some users occasionally encounter a panic during OS hardware detection on 2.4 running under ESXi 6.5.0 U1 (Build 6765664) -- before handoff to our code -- in `vga_bitblt_text()`. Because it is before the handoff to our code, DDB is not yet configured so the VM drops to a `db>` prompt and waits for input. The crash is unusual in that it does not happen to every VM at every boot. It is random and only affects a small number of reboot attempts. The crash happens before disks are mounted so filesystem corruption is not a concern.

This appears to be a confirmed FreeBSD issue:

https://bugs.freebsd.org/bugzilla/show_bug.cgi?id=217282 (has a patch available, and it's in -CURRENT)

https://bugs.freebsd.org/bugzilla/show_bug.cgi?id=220923 (appears to be a duplicate of 217282)

From reading those bug reports, it appears to be a race condition in the VT code.

A possible workaround is to set `debug.debugger_on_panic=0` in `/boot/loader.conf.local` and then configure a tunable in the pfSense GUI to set `debug.debugger_on_panic=1` so that unrelated crash dumps can be collected afterward. That will not stop the panic, but it will allow the VM to reboot itself until it succeeds.

If the crash is in VT, another possible solution would be to switch affected VMs to the `sc` console by setting `kern.vty=sc` in `/boot/loader.conf.local`

So far over 12 VMs on 2.4.x and FreeBSD 11 I have only managed to make it happen once on my lab ESXi host. See the attached image for the backtrace.

History

#1 - 10/11/2017 09:35 AM - Jim Pingle

- Description updated

#2 - 10/11/2017 12:51 PM - Luiz Souza

- Status changed from Confirmed to Feedback

- % Done changed from 0 to 100

The fix is already merge and will be available on next snapshot.

#3 - 10/12/2017 09:18 AM - Jim Pingle

For anyone experiencing this crash in the meantime, adding `kern.vty=sc` to `/boot/loader.conf.local` is confirmed to work around the issue. This can also be added to `/boot/loader.conf.local` before upgrade if someone is worried they may encounter this race condition.

Once a patched version is available in a release, that change will no longer be necessary.

#4 - 10/12/2017 05:47 PM - Gianluca Toso

- File `pfSense_panic.png` added

For information, the same problem occurs in Workstation 12.5.7 (build 5813279), vm hardware version 11.
It happened to me 3 consecutive times and then no more despite several restarts.

#5 - 10/13/2017 08:50 AM - Jim Pingle

For reference, at least one person appears to have encountered it on ESX 5.5 as well, though the majority of users are only seeing it on 6.5.0 U1.

#6 - 10/17/2017 12:14 PM - Jim Pingle

I can't reproduce this on 2.4.1 snapshots but it was so random before that doesn't give me much confidence.

Anyone else experiencing the issue can try upgrading to a 2.4.1 snapshot to see if it still crashes.

#7 - 10/18/2017 03:38 AM - Constantine Kormashev

- File *vm_bug.png* added

- File *vm_bug2.png* added

Tried on 2 different esxi hosts latest 2.4.1 ova rebooted 20 times each VM. Once got error for 2nd VM.

```
da0: 320.000MB/s transfers (160.000MHz, offset 127, 16bit)
da0: Command Queueing enabled
da0: 8192MB (16777216 512 byte sectors)
da0: quirks=0x140<RETRY_BUSY,STRICT_UNMAP>
cd0 at ata1 bus 0 schs1 target 0 lun 0
Trying to mount root from ufs:/dev/gpt/pfSense [rw]...
GEOM: da0: the secondary GPT header is not in the last LBA.

Fatal trap 12: page fault while in kernel mode
cpuid = 0; apic id = 00
fault virtual address = 0x0
fault code           = supervisor read instruction, page not present
instruction pointer   = 0x20:0x0
stack pointer         = 0x28:0xfffffe001c9eb908
frame pointer         = 0x28:0xfffffe001c9eb9b0
code segment          = base 0x0, limit 0xfffff, type 0x1b
                      = DPL 0, pres 1, long 1, def32 0, gran 1
processor eflags      = interrupt enabled, resume, IOPL = 0
current process       = 12 (swi4: clock (0))
[ thread pid 12 tid 100008 ]
Stopped at           0
db> ?
```

File View VM



```
db> run
?
KDB: reentering
KDB: stack backtrace:
db_trace_self_wrapper() at db_trace_self_wrapper+0x2b/frame 0xfffffe001c9eb3a0
kdb_reenter() at kdb_reenter+0x2f/frame 0xfffffe001c9eb3b0
db_run_cmd() at db_run_cmd+0x1a/frame 0xfffffe001c9eb3c0
db_command() at db_command+0x2bf/frame 0xfffffe001c9eb490
db_command_loop() at db_command_loop+0x64/frame 0xfffffe001c9eb4a0
db_trap() at db_trap+0xef/frame 0xfffffe001c9eb530
kdb_trap() at kdb_trap+0x193/frame 0xfffffe001c9eb5c0
trap_fatal() at trap_fatal+0x2e2/frame 0xfffffe001c9eb610
trap_pfault() at trap_pfault+0x49/frame 0xfffffe001c9eb670
trap() at trap+0x286/frame 0xfffffe001c9eb830
calltrap() at calltrap+0x8/frame 0xfffffe001c9eb830
--- trap 0xc, rip = 0, rsp = 0xfffffe001c9eb908, rbp = 0xfffffe001c9eb9b0 ---
??() at 0/frame 0xfffffe001c9eb9b0
softclock() at softclock+0xb9/frame 0xfffffe001c9eb9e0
intr_event_execute_handlers() at intr_event_execute_handlers+0xec/frame 0xfffffe001c9eba20
ithread_loop() at ithread_loop+0xd6/frame 0xfffffe001c9eba70
fork_exit() at fork_exit+0x85/frame 0xfffffe001c9ebab0
fork_trampoline() at fork_trampoline+0xe/frame 0xfffffe001c9ebab0
--- trap 0, rip = 0, rsp = 0, rbp = 0 ---
--More--
```

#8 - 10/18/2017 09:23 AM - Jim Pingle

- File Selection_709.png added

- Status changed from Feedback to Assigned

Ditto, I see a similar crash. I had to reboot 5 VMs a few times before one of them failed.

```
Fatal trap 12: page fault while in kernel mode
cpuid = 0; apic id = 00
fault virtual address   = 0x0
fault code              = supervisor read instruction, page not present
instruction pointer     = 0x20:0x0
stack pointer          = 0x28:0xfffffe0028164a48
frame pointer          = 0x28:0xfffffe0028164af0
code segment           = base 0x0, limit 0xfffff, type 0x1b
                       = DPL 0, pres 1, long 1, def32 0, gran 1
processor eflags       = interrupt enabled, resume, IOPL = 0
current process        = 12 (swi4: clock (0))
[ thread pid 12 tid 100008 ]
Stopped at            0
db> bt
Tracing pid 12 tid 100008 td 0xfffff800031c1560
??() at 0
softclock_call_cc() at softclock_call_cc+0x13e/frame 0xfffffe0028164af0
softclock() at softclock+0xb9/frame 0xfffffe0028164b20
intr_event_execute_handlers() at intr_event_execute_handlers+0xec/frame 0xfffffe0028164b60
ithread_loop() at ithread_loop+0xd6/frame 0xfffffe0028164bb0
fork_exit() at fork_exit+0x85/frame 0xfffffe0028164bf0
fork_trampoline() at fork_trampoline+0xe/frame 0xfffffe0028164bf0
--- trap 0, rip = 0, rsp = 0, rbp = 0 ---
db> █
```

#9 - 10/18/2017 11:32 AM - Luiz Souza

The recent crashes seems unrelated to the original crash in VT.

They actually seem to happen quite late in the kernel boot to be related to a VT crash.

We should open a new issue to track this new crash.

#10 - 10/18/2017 11:36 AM - Luiz Souza

Ok, I see now the two different crashes on the OP post.

While I take back part of what I said before, It still doesn't look related to the VT.

#11 - 10/18/2017 11:50 AM - Jim Pingle

To rule that out we should setup the kern.vty=sc workaround and continue testing for a bit to see if it still crashes. If it does, then it must be something

new.

#12 - 10/18/2017 02:09 PM - Jim Pingle

- Status changed from Assigned to Resolved

I ran some more tests:

kern.vty=sc ADDED to /boot/loader.conf.local: 72 reboots (6 VMs, 12 reboots each), no crashes
kern.vty=sc REMOVED from /boot/loader.conf.local: 72 reboots (6 VMs, 12 reboots each), no crashes

So that's 144 crash-free reboots total on 2.4.1, and half of those should have met the conditions to trigger the VT race if it was still a problem.

I was hoping to reproduce it again to see if I was related, but now I'm not seeing it either way.

If we can manage to reproduce the conditions for that swi/clock crash we can open a new ticket for it.

#13 - 10/19/2017 01:30 AM - Nicolas Liaudat

Jim Pingle wrote:

For reference, at least one person appears to have encountered it on ESX 5.5 as well, though the majority of users are only seeing it on 6.5.0 U1.

Problem confirmed on esxi 6.0

Files

Selection_704.png	15.6 KB	10/11/2017	Jim Pingle
pfsense_panic.png	332 KB	10/12/2017	Gianluca Toso
vm_bug.png	86.9 KB	10/18/2017	Constantine Kormashev
vm_bug2.png	122 KB	10/18/2017	Constantine Kormashev
Selection_709.png	14.6 KB	10/18/2017	Jim Pingle