Bug 208767 - kernel stack overflow due to Lazy update IOAPIC on an x86_64 *host*, when gpu is passthrough to macos guest vm

Status: REOPENED

Alias: None

Product: Virtualization
Component: kvm (show other bugs)
Hardware: All Linux

Importance: P1 normal
Assignee: virtualization_kvm

URL:
Keywords:

Depends on:
Blocks:

Reported: 2020-08-02 09:01 UTC by Yani Stoyanov
Modified: 2022-04-29 03:14 UTC (History)
CC List: 6 users (show)

See Also:
Kernel Version: 5.6 up to and including 5.7.11
Tree: Mainline
Regression: Yes

Attachments

Add an attachment (proposed patch, testcase, etc.)

Yani Stoyanov **2020-08-02 09:01:40 UTC** **Description**

I have fedora 32 host with latest kernel on a double xeon v5 2630 workstation asus board and few vm with assigned gpus to them (linux windows and macos).

I notice that after I think kernel 5.5.19, more concrete the introduction of:

ioapic_lazy_update_eoi(ioapic, irq); in: ioapic.c my macos guest stop booting when there is a gpu assigned to them. I have old gforce 970 and rx470, I try with each of them the result was always the same. I also try with mac os Sierra and Catalina (for catalina only amd gpu us supported) and again the vm hangs. After the hang the whole lib virt service became not responsible.

I test this with only the gpu assigned to the vm to exclude cases with multiple devices, I also try different cpu core configurations and the result was the same.

I try to comment the mentioned:

```
if (edge && kvm_apicv_activated(ioapic->kvm))
ioapic_lazy_update_eoi(ioapic, irq);
```

in ioapic.c, rebuild the kernel from source and try with my custom one and then the macos vm start correctly.

When the issue appear i notice this ind the dmesg output:

```
5533.660264] BUG: stack guard page was hit at 0000000072715902 (stack is
0000000078c6c553..000000008fa11e86)
[ 5533.660273] kernel stack overflow (double-fault): 0000 [#1] SMP PTI
[ 5533.660277] CPU: 10 PID: 6476 Comm: qemu-system-x86 Not tainted 5.7.10-
201.fc32.x86_64 #1
[ 5533.660279] Hardware name: ASUSTeK COMPUTER INC. Z10PE-D16 WS/Z10PE-D16 WS, BIOS
4101 06/12/2019
[ 5533.660323] RIP: 0010:kvm_set_irq+0x20/0x130 [kvm]
[ 5533.660327] Code: c3 66 0f 1f 84 00 00 00 00 00 0f 1f 44 00 00 41 57 41 89 d7 41
56 41 55 41 54 49 89 fc 55 89 cd 53 89 f3 48 81 ec c8 00 00 00 <44> 89 44 24 04 0f
1f 44 00 00 4d 8d ac 24 c8 17 02 00 4c 89 ef e8
[ 5533.660329] RSP: 0018:ffffadba89cc7f70 EFLAGS: 00010282
[ 5533.660332] RAX: ffffffffc06f7da0 RBX: 0000000000000001 RCX: 0000000000000000
[ 5533.660334] RDX: 000000000000000b RSI: 0000000000000001 RDI: ffffadba8aa31000
[ 5533.660335] RBP: 0000000000000000 R08: 0000000000000000 R09: 0000000000000000
[ 5533.660337] R10: 0000000000000000 R11: ffff9b80e7498000 R12: ffffadba8aa31000
[ 5533.660338] R13: 0000000000000000 R14: ffffadba8aa527c8 R15: 000000000000000b
[ 5533.660341] FS: 00007f0355dd0ec0(0000) GS:ffff9b92bfa80000(0000)
knlGS:0000000000000000
[ 5533.660343] CS: 0010 DS: 0000 ES: 0000 CR0: 0000000080050033
[ 5533.660344] CR2: ffffadba89cc7f68 CR3: 0000000fcb4aa001 CR4: 00000000001626a0
[ 5533.660346] Call Trace:
[ 5533.660373] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.660391] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.660415] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.660437] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.660458] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.660476] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.660497] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.660516] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.660534] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.660551] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.660571] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.660589] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.660607] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.660625] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.660644] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.660662] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.660680] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.660697] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.660716] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.660735] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.660753] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.660771] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.660789] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.660807] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.660825] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.660858] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.660877] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.660895] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.660913] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.660931] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.660949] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.660966] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.660984] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.661000] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.661019] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.661037] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.661055] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.661072] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.661090] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.661108] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.661125] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.661142] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.661160] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.661178] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.661196] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.661213] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.661231] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.661248] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.661266] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.661282] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.661301] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.661318] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.661336] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.661353] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.661371] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.661389] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.661407] irqfd_resampler_ack+0x32/0x90 [kvm]
```

```
[ 5533.661423] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.661441] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.661459] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.661477] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.661495] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.661513] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.661530] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.661548] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.661564] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.661583] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.661600] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.661618] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.661635] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.661653] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.661671] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.661688] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.661705] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.661723] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.661741] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.661759] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.661776] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.661794] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.661811] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.661829] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.661845] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.661863] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.661881] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.661899] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.661916] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.661934] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.661951] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.661969] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.661985] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662003] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.662021] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.662039] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.662056] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.662074] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.662092] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.662109] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.662125] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662144] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.662161] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.662179] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.662196] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.662214] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.662232] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.662249] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.662266] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662284] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.662302] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.662319] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.662336] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.662354] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.662372] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.662389] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.662406] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662424] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.662442] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.662459] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.662477] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.662494] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.662512] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.662529] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.662546] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662564] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.662582] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.662600] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.662617] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.662635] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.662652] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.662670] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.662687] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662704] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.662722] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.662740] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.662757] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.662775] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.662792] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.662809] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.662826] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662844] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.662861] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.662879] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.662896] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.662914] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.662931] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.662948] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.662965] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.662983] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663001] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.663019] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.663036] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.663054] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.663071] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.663088] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.663105] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.663123] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663140] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.663158] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.663175] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.663193] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.663210] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.663228] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.663244] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.663262] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663279] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.663297] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.663314] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.663332] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.663349] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.663367] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.663384] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.663401] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663419] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.663436] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.663453] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.663471] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.663488] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.663506] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.663522] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.663540] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663557] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.663575] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.663592] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.663610] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.663627] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.663645] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.663661] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.663679] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663697] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.663715] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.663732] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.663749] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.663766] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.663784] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.663801] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.663818] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663836] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.663854] kvm_ioapic_set_irq+0x61/0x90 [kvm]
```

```
[ 5533.663871] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.663888] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.663906] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.663912] ? x86_configure_nx+0x40/0x40
[ 5533.663917] ? cpumask_next+0x17/0x20
[ 5533.663934] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.663951] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.663969] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.663986] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.664004] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.664021] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.664039] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.664056] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.664074] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.664090] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.664108] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.664125] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.664143] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.664160] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.664178] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.664195] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.664202] ? __switch_to_asm+0x34/0x70
[ 5533.664205] ? __switch_to_asm+0x40/0x70
[ 5533.664208] ? __switch_to_asm+0x34/0x70
[ 5533.664213] ? __switch_to_xtra+0x10a/0x500
[ 5533.664230] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.664246] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.664264] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.664281] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.664299] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.664316] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.664333] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.664351] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.664371] ? kvm_irq_delivery_to_apic_fast+0xd8/0x130 [kvm]
[ 5533.664389] irqfd_resampler_ack+0x32/0x90 [kvm]
[ 5533.664405] kvm_notify_acked_irq+0xc4/0xe0 [kvm]
[ 5533.664424] kvm_ioapic_update_eoi_one.isra.0+0x3c/0x130 [kvm]
[ 5533.664442] ioapic_set_irq+0x21d/0x260 [kvm]
[ 5533.664460] kvm_ioapic_set_irq+0x61/0x90 [kvm]
[ 5533.664477] kvm_set_irq+0xa5/0x130 [kvm]
[ 5533.664495] ? kvm_hv_set_sint+0x20/0x20 [kvm]
[ 5533.664513] ? kvm_set_ioapic_irq+0x20/0x20 [kvm]
[ 5533.664518] ? get_order+0x20/0x20
[ 5533.664538] kvm_vm_ioctl_irq_line+0x23/0x30 [kvm]
[ 5533.664556] kvm_vm_ioctl+0x163/0xd40 [kvm]
[ 5533.664559] ? get_order+0x20/0x20
[ 5533.664564] ? recalibrate_cpu_khz+0x10/0x10
[ 5533.664566] ? poll_select_finish+0x15b/0x1f0
[ 5533.664569] ksys_ioctl+0x82/0xc0
[ 5533.664573] __x64_sys_ioctl+0x16/0x20
[ 5533.664577] do_syscall_64+0x5b/0xf0
[ 5533.664582] entry_SYSCALL_64_after_hwframe+0x44/0xa9
[ 5533.664585] RIP: 0033:0x7f03576993bb
[ 5533.664588] Code: 0f 1e fa 48 8b 05 dd aa 0c 00 64 c7 00 26 00 00 00 48 c7 c0 ff
ff ff ff c3 66 0f 1f 44 00 00 f3 0f 1e fa b8 10 00 00 00 0f 05 <48> 3d 01 f0 ff ff
73 01 c3 48 8b 0d ad aa 0c 00 f7 d8 64 89 01 48
[ 5533.664590] RSP: 002b:00007fffe4ba8598 EFLAGS: 00000246 ORIG_RAX:
0000000000000010
[ 5533.664593] RAX: ffffffffffffffda RBX: 000055719fbc23d0 RCX: 00007f03576993bb
[ 5533.664595] RDX: 00007fffe4ba8600 RSI: ffffffffc008ae67 RDI: 000000000000000e
[ 5533.664596] RBP: 000000000000000b R08: 0000000000000000 R09: 000000000000001c
[ 5533.664598] R10: 00007fffe4ba85b0 R11: 0000000000000246 R12: 0000000000000001
[ 5533.664599] R13: 00007fffe4ba8644 R14: 000000000000009f R15: 0000000000000000
[ 5533.664603] Modules linked in: xt_CHECKSUM xt_MASQUERADE xt_conntrack ipt_REJECT
nf_nat_tftp nf_conntrack_tftp tun bridge stp llc nft_objref nf_conntrack_netbios_ns
nf_conntrack_broadcast nft_fib_inet nft_fib_ipv4 nft_fib_ipv6 nft_fib
nft_reject_inet nf_reject_ipv4 nf_reject_ipv6 nft_reject nft_ct nft_chain_nat
nf_tables ebtable_nat ebtable_broute ip6table_nat ip6table_mangle ip6table_raw
ip6table_security iptable_nat nf_nat nf_conntrack nf_defrag_ipv6 nf_defrag_ipv4
libcrc32c iptable_mangle iptable_raw iptable_security ip_set nfnetlink
ebtable_filter ebtables ip6table_filter ip6_tables iptable_filter sunrpc
intel_rapl_msr intel_rapl_common sb_edac x86_pkg_temp_thermal coretemp kvm_intel
kvm eeepc_wmi rapl asus_wmi sparse_keymap intel_cstate rfkill mei_me ipmi_ssif
intel_uncore video wmi_bmof pcspkr joydev i2c_i801 mei lpc_ich ipmi_si ipmi_devintf
ipmi_msghandler acpi_power_meter ip_tables hid_logitech_hidpp igbvf ast
drm_vram_helper drm_ttm_helper ttm drm_kms_helper crct10dif_pclmul
[ 5533.664645] crc32_pclmul crc32c_intel drm ghash_clmulni_intel mxm_wmi igb
hid_logitech_dj dca i2c_algo_bit wmi vfio_pci irqbypass vfio_virqfd
vfio_iommu_type1 vfio fuse
[ 5533.664663] ---[ end trace 662e3e16e1eb8bcc ]---
[ 5533.680637] RIP: 0010:kvm_set_irq+0x20/0x130 [kvm]
[ 5533.680641] Code: c3 66 0f 1f 84 00 00 00 00 00 0f 1f 44 00 00 41 57 41 89 d7 41
56 41 55 41 54 49 89 fc 55 89 cd 53 89 f3 48 81 ec c8 00 00 00 <44> 89 44 24 04 0f
1f 44 00 00 4d 8d ac 24 c8 17 02 00 4c 89 ef e8
[ 5533.680643] RSP: 0018:ffffadba89cc7f70 EFLAGS: 00010282
[ 5533.680645] RAX: ffffffffc06f7da0 RBX: 0000000000000001 RCX: 0000000000000000
[ 5533.680647] RDX: 000000000000000b RSI: 0000000000000001 RDI: ffffadba8aa31000
[ 5533.680648] RBP: 0000000000000000 R08: 0000000000000000 R09: 0000000000000000
[ 5533.680650] R10: 0000000000000000 R11: ffff9b80e7498000 R12: ffffadba8aa31000
[ 5533.680651] R13: 0000000000000000 R14: ffffadba8aa527c8 R15: 000000000000000b
[ 5533.680654] FS: 00007f0355dd0ec0(0000) GS:ffff9b92bfa80000(0000)
knlGS:0000000000000000
[ 5533.680655] CS: 0010 DS: 0000 ES: 0000 CR0: 0000000080050033
[ 5533.680657] CR2: ffffadba89cc7f68 CR3: 0000000fcb4aa001 CR4: 00000000001626a0
[ 5533.680870] ------------[ cut here ]------------
[ 5533.680877] WARNING: CPU: 10 PID: 0 at kernel/rcu/tree.c:569
rcu_eqs_enter.constprop.0+0xb4/0xc0
[ 5533.680878] Modules linked in: xt_CHECKSUM xt_MASQUERADE xt_conntrack ipt_REJECT
nf_nat_tftp nf_conntrack_tftp tun bridge stp llc nft_objref nf_conntrack_netbios_ns
nf_conntrack_broadcast nft_fib_inet nft_fib_ipv4 nft_fib_ipv6 nft_fib
nft_reject_inet nf_reject_ipv4 nf_reject_ipv6 nft_reject nft_ct nft_chain_nat
nf_tables ebtable_nat ebtable_broute ip6table_nat ip6table_mangle ip6table_raw
ip6table_security iptable_nat nf_nat nf_conntrack nf_defrag_ipv6 nf_defrag_ipv4
libcrc32c iptable_mangle iptable_raw iptable_security ip_set nfnetlink
ebtable_filter ebtables ip6table_filter ip6_tables iptable_filter sunrpc
intel_rapl_msr intel_rapl_common sb_edac x86_pkg_temp_thermal coretemp kvm_intel
kvm eeepc_wmi rapl asus_wmi sparse_keymap intel_cstate rfkill mei_me ipmi_ssif
intel_uncore video wmi_bmof pcspkr joydev i2c_i801 mei lpc_ich ipmi_si ipmi_devintf
ipmi_msghandler acpi_power_meter ip_tables hid_logitech_hidpp igbvf ast
drm_vram_helper drm_ttm_helper ttm drm_kms_helper crct10dif_pclmul
[ 5533.680901] crc32_pclmul crc32c_intel drm ghash_clmulni_intel mxm_wmi igb
hid_logitech_dj dca i2c_algo_bit wmi vfio_pci irqbypass vfio_virqfd
vfio_iommu_type1 vfio fuse
[ 5533.680910] CPU: 10 PID: 0 Comm: swapper/10 Tainted: G D 5.7.10-201.fc32.x86_64
#1
[ 5533.680912] Hardware name: ASUSTeK COMPUTER INC. Z10PE-D16 WS/Z10PE-D16 WS, BIOS
4101 06/12/2019
[ 5533.680915] RIP: 0010:rcu_eqs_enter.constprop.0+0xb4/0xc0
[ 5533.680917] Code: 0f c1 83 d8 00 00 00 65 8b 15 78 73 eb 71 65 48 8b 04 25 c0 8b
01 00 89 90 7c 07 00 00 5b 5d c3 48 89 ef e8 9e fe ff ff eb bd <0f> 0b e9 75 ff ff
ff 0f 1f 44 00 00 0f 1f 44 00 00 41 57 41 56 41
[ 5533.680919] RSP: 0018:ffffadba8635bee0 EFLAGS: 00010002
[ 5533.680921] RAX: ffff9b92bfaabc00 RBX: 0000000000002bc00 RCX: 0000000000000000
[ 5533.680922] RDX: 4000000000000000 RSI: ffffadba8635bea0 RDI: 0000050867d45ea5
[ 5533.680924] RBP: 000000000000000a R08: 0000000000000001 R09: 0000000000000020
[ 5533.680925] R10: 0000050867d45ea5 R11: 0000000000000004 R12: 0000000000000000
[ 5533.680927] R13: 0000000000000000 R14: 0000000000000000 R15: 0000000000000000
[ 5533.680929] FS: 0000000000000000(0000) GS:ffff9b92bfa80000(0000)
knlGS:0000000000000000
[ 5533.680930] CS: 0010 DS: 0000 ES: 0000 CR0: 0000000080050033
[ 5533.680932] CR2: ffffadba89cc7f68 CR3: 0000000fcb4aa001 CR4: 00000000001626a0
[ 5533.680933] Call Trace:
[ 5533.680939] do_idle+0x1c6/0x240
[ 5533.680943] cpu_startup_entry+0x19/0x20
[ 5533.680950] start_secondary+0x15c/0x1a0
[ 5533.680954] secondary_startup_64+0xb6/0xc0
[ 5533.680957] ---[ end trace 662e3e16e1eb8bcd ]---
[ 5549.993848] kvm [3589]: vcpu2, guest rIP: 0x7f04196453d5 ignored rdmsr: 0x122
[ 5549.994176] kvm [3589]: vcpu3, guest rIP: 0x7ff35028d3d5 ignored rdmsr: 0x122
[ 5549.994451] kvm [3589]: vcpu0, guest rIP: 0x7f274147e3d5 ignored rdmsr: 0x122
```

```
[ 5550.004411] kvm [3589]: vcpu2, guest rIP: 0x7f04191877cb ignored rdmsr: 0x122
[ 5550.004587] kvm [3589]: vcpu2, guest rIP: 0x7f0418db732a ignored rdmsr: 0x122
[ 5550.004796] kvm [3589]: vcpu0, guest rIP: 0x7f2740fc07cb ignored rdmsr: 0x122
[ 5550.004802] kvm [3589]: vcpu3, guest rIP: 0x7ff34fdcf7cb ignored rdmsr: 0x122
[ 5550.005016] kvm [3589]: vcpu0, guest rIP: 0x7f2740bf032a ignored rdmsr: 0x122
[ 5550.005023] kvm [3589]: vcpu3, guest rIP: 0x7ff34f9ff32a ignored rdmsr: 0x122
[ 5587.446873] kvm [3683]: vcpu0, guest rIP: 0x80061bfd8 ignored rdmsr: 0x122
[ 5587.463361] kvm [3683]: vcpu1, guest rIP: 0x80061bfd8 ignored rdmsr: 0x122
[ 5587.607803] kvm [3683]: vcpu2, guest rIP: 0x800606fd8 ignored rdmsr: 0x122
[ 5587.623100] kvm [3683]: vcpu0, guest rIP: 0x800629fd8 ignored rdmsr: 0x122
[ 5587.626789] kvm [3683]: vcpu3, guest rIP: 0x800603fd8 ignored rdmsr: 0x122
[ 5587.629251] kvm [3683]: vcpu0, guest rIP: 0x800604fd8 ignored rdmsr: 0x122
[ 5587.629969] kvm [3683]: vcpu0, guest rIP: 0x800629fd8 ignored rdmsr: 0x122
[ 5587.631970] kvm [3683]: vcpu1, guest rIP: 0x800603fd8 ignored rdmsr: 0x122
[ 5587.635258] kvm [3683]: vcpu3, guest rIP: 0x80060afd8 ignored rdmsr: 0x122
[ 5617.847070] kvm [3683]: vcpu3, guest rIP: 0x8006affd8 ignored rdmsr: 0x122
[ 5617.847548] kvm [3683]: vcpu3, guest rIP: 0x8018c739a ignored rdmsr: 0x122
[ 5617.866134] kvm [3683]: vcpu0, guest rIP: 0x8006affd8 ignored rdmsr: 0x122
[ 5617.866437] kvm [3683]: vcpu0, guest rIP: 0x8018c739a ignored rdmsr: 0x122
[ 5617.884771] kvm [3683]: vcpu2, guest rIP: 0x8006affd8 ignored rdmsr: 0x122
[ 5617.885059] kvm [3683]: vcpu2, guest rIP: 0x8018c739a ignored rdmsr: 0x122
[ 5617.897966] kvm [3683]: vcpu1, guest rIP: 0x8006affd8 ignored rdmsr: 0x122
[ 5617.898493] kvm [3683]: vcpu1, guest rIP: 0x8018c739a ignored rdmsr: 0x122
[ 5647.472842] kvm [3683]: vcpu3, guest rIP: 0x80061bfd8 ignored rdmsr: 0x122
[ 5647.493155] kvm [3683]: vcpu0, guest rIP: 0x80061bfd8 ignored rdmsr: 0x122
[ 5647.801439] kvm [3683]: vcpu1, guest rIP: 0x800606fd8 ignored rdmsr: 0x122
[ 5647.821566] kvm [3683]: vcpu1, guest rIP: 0x800629fd8 ignored rdmsr: 0x122
[ 5647.824748] kvm [3683]: vcpu2, guest rIP: 0x800603fd8 ignored rdmsr: 0x122
[ 5647.826885] kvm [3683]: vcpu1, guest rIP: 0x800604fd8 ignored rdmsr: 0x122
[ 5647.827646] kvm [3683]: vcpu1, guest rIP: 0x800629fd8 ignored rdmsr: 0x122
[ 5647.830756] kvm [3683]: vcpu3, guest rIP: 0x800603fd8 ignored rdmsr: 0x122
[ 5647.835779] kvm [3683]: vcpu2, guest rIP: 0x80060afd8 ignored rdmsr: 0x122
[ 5648.046010] kvm [3683]: vcpu2, guest rIP: 0x800629fd8 ignored rdmsr: 0x122
[ 5666.327133] kvm_get_msr_common: 21 callbacks suppressed
[ 5666.327135] kvm [4004]: vcpu2, guest rIP: 0x7ffe989ea0c7 ignored rdmsr: 0x122
[ 5666.329557] kvm [4004]: vcpu3, guest rIP: 0x7ff6c088934b ignored rdmsr: 0x122
[ 5666.333302] kvm [4004]: vcpu0, guest rIP: 0x7ffec2ae6807 ignored rdmsr: 0x122
[ 5666.333326] kvm [4004]: vcpu0, guest rIP: 0x7ffec2ae6807 ignored rdmsr: 0x122
[ 5666.333478] kvm [4004]: vcpu0, guest rIP: 0x7ffec2bbe257 ignored rdmsr: 0x122
[ 5666.333804] kvm [4004]: vcpu0, guest rIP: 0x7ffec38e13bf ignored rdmsr: 0x122
[ 5666.333928] kvm [4004]: vcpu3, guest rIP: 0x7ffec38e13bf ignored rdmsr: 0x122
[ 5666.337446] kvm [4004]: vcpu3, guest rIP: 0x7ff6c07d8344 ignored rdmsr: 0x122
[ 5666.375209] kvm [4004]: vcpu1, guest rIP: 0x7ffec2ae6807 ignored rdmsr: 0x122
[ 5666.375235] kvm [4004]: vcpu1, guest rIP: 0x7ffec2ae6807 ignored rdmsr: 0x122
[ 5684.258354] kvm_get_msr_common: 71 callbacks suppressed
[ 5684.258356] kvm [4004]: vcpu1, guest rIP: 0x7ffec2ae6807 ignored rdmsr: 0x122
[ 5684.258386] kvm [4004]: vcpu1, guest rIP: 0x7ffec2ae6807 ignored rdmsr: 0x122
[ 5684.258551] kvm [4004]: vcpu1, guest rIP: 0x7ffec2bbe257 ignored rdmsr: 0x122
[ 5684.258739] kvm [4004]: vcpu1, guest rIP: 0x7ffec38e13bf ignored rdmsr: 0x122
[ 5684.258772] kvm [4004]: vcpu1, guest rIP: 0x7ffec38e13bf ignored rdmsr: 0x122
[ 5684.264699] kvm [4004]: vcpu6, guest rIP: 0x7ffec2e83e8f ignored rdmsr: 0x122
[ 5684.264731] kvm [4004]: vcpu6, guest rIP: 0x7ffec2e83e8f ignored rdmsr: 0x122
[ 5684.264985] kvm [4004]: vcpu6, guest rIP: 0x7ffec09dd827 ignored rdmsr: 0x122
[ 5684.265019] kvm [4004]: vcpu6, guest rIP: 0x7ffec09dd827 ignored rdmsr: 0x122
[ 5684.266914] kvm [4004]: vcpu6, guest rIP: 0x7ffebde4b43b ignored rdmsr: 0x12
```

I am not a linux power user so in case of needing more information please provide detailed steps how can I generate it so I can be more useful.

Paolo Bonzini 2020-08-02 09:19:16 UTC Comment 1

This should have been fixed by commit 8be8f932e3db5fe4ed178b8892eeffeab530273a in Linux 5.7.

Yani Stoyanov 2020-08-02 10:36:19 UTC Comment 2

I was thinking the same thing when I saw: ~~https://bugzilla.kernel.org/show_bug.cgi?id=207489~~

I write a comment there but start realizing that the reason for my issue may be something different since it happens only with my macos vm-s. Currently I am using kernel-5.7.10-201.fc32.x86_64 which should include the patch.

And I mentioned bug people are complaining about windows guests. I have 2 windows 10 machines and they are working fine no issues there the problem appear only on my macos vm.

Yani Stoyanov 2020-08-02 17:45:43 UTC Comment 3

(In reply to Paolo Bonzini from comment #1)
> This should have been fixed by commit
> 8be8f932e3db5fe4ed178b8892eeffeab530273a in Linux 5.7.

This commit is already merged to kernel-5.7.10-201.fc32.x86_64 right?

Jim Mattson 2020-08-03 20:39:21 UTC Comment 4

On Sun, Aug 2, 2020 at 2:01 AM <bugzilla-daemon@bugzilla.kernel.org> wrote:
>
> https://bugzilla.kernel.org/show_bug.cgi?id=208767

>
>   Bug ID: 208767
>   Summary: kernel stack overflow due to Lazy update IOAPIC on an
>   x86_64 *host*, when gpu is passthrough to macos guest
>   vm
>   Product: Virtualization
>   Version: unspecified
>   Kernel Version: 5.6 up to and including 5.7
>   Hardware: All
>   OS: Linux
>   Tree: Mainline
>   Status: NEW
>   Severity: normal
>   Priority: P1
>   Component: kvm
>   Assignee: virtualization_kvm@kernel-bugs.osdl.org
>   Reporter: yaweb@mail.bg
>   Regression: No

>
> I have fedora 32 host with latest kernel on a double xeon v5 2630 workstation
> asus board and few vm with assigned gpus to them (linux windows and macos).

I didn't think the Mac OS X license agreement permitted running it on non-Apple hardware. Has this changed?

Yani Stoyanov 2020-08-04 00:25:20 UTC Comment 5

(In reply to Jim Mattson from comment #4)
> On Sun, Aug 2, 2020 at 2:01 AM <bugzilla-daemon@bugzilla.kernel.org> wrote:
> >
> > https://bugzilla.kernel.org/show_bug.cgi?id=208767
> >
> >   Bug ID: 208767
> >   Summary: kernel stack overflow due to Lazy update IOAPIC on an
> >   x86_64 *host*, when gpu is passthrough to macos guest
> >   vm
> >   Product: Virtualization
> >   Version: unspecified

> >   Kernel Version: 5.6 up to and including 5.7
> >   Hardware: All
> >   OS: Linux
> >   Tree: Mainline
> >   Status: NEW
> >   Severity: normal
> >   Priority: P1
> >   Component: kvm
> >   Assignee: virtualization_kvm@kernel-bugs.osdl.org
> >   Reporter: vaweb@mail.bg
> >   Regression: No
> >
> > I have fedora 32 host with latest kernel on a double xeon v5 2630
> workstation
> > asus board and few vm with assigned gpus to them (linux windows and macos).
>
> I didn't think the Mac OS X license agreement permitted running it on
> non-Apple hardware. Has this changed?

Jim Mattson, I guess official it is not support by as I wrote in the description of the issue the problem is in the mentioned function. I tested it and if comment the lines

```
if (edge && kvm_apicv_activated(ioapic->kvm))
ioapic_lazy_update_eoi(ioapic, irq);
```

It boots fine, if the function invocation is not commented I kernel stack overflow so the bug is for it it should not matter what case it right?

Yani Stoyanov 2020-08-06 09:09:28 UTC Comment 6

I am not sure if this is relevant but there was old bug which explains how osx configure IOAPIC with the wrong polarity bit values. I may be interesting to take a look (I know it is from 6 years ago).

https://www.contrib.andrew.cmu.edu/~somlo/OSXKVM/index_old.html

the part:

ACPI-compliant operating systems are expected to query the firmware for an indication of which polarity type (ActiveLow or ActiveHigh) to use for any devices with level-triggered interrupts, and to configure the IOAPIC registers accordingly. Both QEMU and KVM have accumulated a significant number of optimizations based on the assumption that guest operating systems use ActiveHigh polarity, and are coded to assume that "physical" and "logical" IRQ line states are in sync. Even when a misbehaving guest OS (you guessed it, OS X does this) ignores the ACPI polarity hint (which in QEMU/KVM is ActiveLow, i.e. "physical"=="logical") and configures the virtual IOAPIC with the wrong polarity bit values, both QEMU and KVM will mostly use "logical" IRQ line levels.

Alex Williamson 2020-10-07 22:45:49 UTC Comment 7

(In reply to Paolo Bonzini from comment #1)
> This should have been fixed by commit
> 8be8f932e3db5fe4ed178b8892eeffeab530273a in Linux 5.7.

This is not fixed and it's not unique to a macos VM, a Linux guest can also reproduce this. I've seen this both during PXE boot and during shutdown with certain NIC combinations (see rhbz1867373). The only workaround is to disable acpiv (kvm_intel.enable_apicv=0). Any suggestions, Paolo?

shantur 2021-02-03 20:32:30 UTC Comment 8

This bug is reproducible on Apple hardware too.

I tried this on MacPro 2013 running QEMU KVM with GPU passthrough and all worked well until the commit with ioapic_lazy_update_eoi came in.