Home › Security

CVE-2022-36123 – A vulnerability in Linux kernel mainline v5.18-rc1 through v5.19-rc6 does not clear statically allocated variables in the block starting symbol (.bss) due to a failed early_xen_iret_patch leading to an

asm_exc_page_fault, or arbitrary code execution

by Sick Codes — July 29, 2022 in Security 💬 11

Title

A vulnerability in Linux kernel mainline v5.18-rc1 through v5.19-rc6 does not clear statically allocated variables in the block starting symbol (.bss) due to a failed early_xen_iret_patch leading to an asm_exc_page_fault, or arbitrary code execution

CVE ID

CVE-2022-36123

CVSS Score

N/A

Internal IDs

Vendor

Kernel.org

Product

Linux Kernel

Product Versions

Kernel v5.18-rc1 through v5.19-rc6

Vulnerability Details

A vulnerability in Linux kernel mainline v5.18-rc1 through v5.19-rc6 may not clear the block starting symbol (.bss) where statically allocated variables in the .bss, affecting XenPV guests, leading to an asm_exc_page_fault, or arbitrary code execution. An unprivileged local attacker on the host, or guest, may potentially use this flaw to cause a NULL Pointer Dereference, kernel oops or denial of service as this allows virtualized devices connected to the Xen IOMMU via xen_set_restricted_virtio_memory_access to potentially access restricted memory. In addition, if kexec is used, the 2nd kernel .bss may contain uninitialized resources and may not be clear.

Vendor Response

Fixed in kernel mainline. clear_bss() now clears the .brk at early boot and xen_set_restricted_virtio_memory_access was added via CONFIG_XEN_VIRTIO kernel config.

Fix: 36e2f161fb01795722f2ff1a24d95f08100333dd

Upstream: 38fa5479b41376dc9d7f57e71c83514285a25ca0

Fixed in 5.18.13 stable
https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.18.13

Fixed in 5.15.56 longterm 5.15.x

https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.15.56

Fixed in 5.10.132 longterm 5.10.x

https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.10.132

Fixed in 5.4.207 longterm 5.4.x

https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.4.207

Fixed in 4.19.253 longterm 4.19.x

https://cdn.kernel.org/pub/linux/kernel/v4.x/ChangeLog-4.19.253

Fixed in 4.14.289 longterm 4.14.x

https://cdn.kernel.org/pub/linux/kernel/v4.x/ChangeLog-4.14.289

Fixed in 4.9.324 longterm 4.9.324

https://cdn.kernel.org/pub/linux/kernel/v4.x/ChangeLog-4.9.324

Proof of Concept

A note was left when v5.18-rc1 was released describing the use of
xen_start_info arch/x86/xen/mmu_pv.c: "The xen_start_info has been taken
care of already in xen_setup_kernel_pagetable"

https://github.com/torvalds/linux/blob/babf0bb978e3c9fce6c4eba6b744c8754fd
43d8e/arch/x86/xen/mmu_pv.c#L1151

After early_xen_iret_patch was added, it was later urgently removed &
cleaned up in x86_urgent_for_v5.19_rc6:

  Merge tag 'x86_urgent_for_v5.19_rc6' of git://git.kernel.org/pub/scm/…

 …linux/kernel/git/tip/tip

 Pull x86 fixes from Borislav Petkov:

  - Prepare for and clear .brk early in order to address XenPV guests
    failures where the hypervisor verifies page tables and uninitialized
    data in that range leads to bogus failures in those checks

- Add any potential setup_data entries supplied at boot to the identity
    pagetable mappings to prevent kexec kernel boot failures. Usually,
    this is not a problem for the normal kernel as those mappings are
    part of the initially mapped 2M pages but if kexec gets to allocate
    the second kernel somewhere else, those setup_data entries need to be
    mapped there too.

  - Fix objtool not to discard text references from the __tracepoints
    section so that ENDBR validation still works

  - Correct the setup_data types limit as it is user-visible, before 5.19
    releases

 * tag 'x86_urgent_for_v5.19_rc6' of git://git.kernel.org/pub/scm/linux/kernel/git/tip/tip:
   x86/boot: Fix the setup data types max limit
   x86/ibt, objtool: Don't discard text references from tracepoint section
   x86/compressed/64: Add identity mappings for setup_data entries
   x86: Fix .brk attribute in linker script
   x86: Clear .brk area at early boot
   x86/xen: Use clear_bss() for Xen PV guests

If clear_bss() is not added, and a user runs kexec, the "normal kernel…
mappings are
part of the initially mapped 2M pages but kexec gets to allocate the
second kernel somewhere else, those setup_data entries need to be mapped
there too."

 * tag 'x86_urgent_for_v5.19_rc6' of git://git.kernel.org/pub/scm/linux/kernel/git/tip/tip:
   x86/boot: Fix the setup data types max limit
   x86/ibt, objtool: Don't discard text references from tracepoint section
   x86/compressed/64: Add identity mappings for setup_data entries
   x86: Fix .brk attribute in linker script
   x86: Clear .brk area at early boot
   x86/xen: Use clear_bss() for Xen PV guests

Virtualized devices connected to the Xen IOMMU can potentially access
restricted memory.

```
[12002.517482] BUG: kernel NULL pointer dereference, address: 0000000000000344
[12002.517487] #PF: supervisor write access in kernel mode
[12002.517489] #PF: error_code(0x0002) - not-present page
[12002.517490] PGD 0 P4D 0
[12002.517493] Oops: 0002 [#1] PREEMPT SMP NOPTI
[12002.517499] RIP: 0010:copy_fpstate_to_sigframe+0xad/0x330
[12002.517505] Code: 1f 44 00 00 48 8d bd 00 02 00 00 be 40 00 00 00 e8 b8 eb 58 00 48 85
c0 75 bc 65 48 8b 1c 25 c0 0b 02 00 65 81 05 9f 73 1e 5d <00> 02 00 00 48 8b 03 f6 c4 40
0f 85 ab 00 00 00 83 83 f8 1a 00 00
[12002.517506] RSP: 0000:ffffb298c2fb3df0 EFLAGS: 00050212
[12002.517508] RAX: 000000005d1e747a RBX: ffffb298c2fb3f58 RCX: 0000000000000008
[12002.517510] RDX: 0000000000000344 RSI: 0000000000000040 RDI: 00007fdb33e8ee80
[12002.517511] RBP: 00007fdb33e8ec80 R08: ffff9510262ef640 R09: 0000000000000000
[12002.517512] R10: 000000000000000b R11: 000000000000000a R12: ffff950c8fca5d40
[12002.517513] R13: ffff950c8fca4080 R14: ffff950c8fca4080 R15: 00007fdb33e8ec80
[12002.517515] FS:  00007fdb32000340(0000) GS:ffff95128f600000(0000)
knlGS:0000000000000000
[12002.517516] CS:  0010 DS: 0000 ES: 0000 CR0: 0000000080050033
[12002.517517] CR2: 0000000000000344 CR3: 0000000204082000 CR4: 0000000000350ef0
[12002.517520] Call Trace:
[12002.517521]  <TASK>
[12002.517522]  ? copy_fpstate_to_sigframe+0x98/0x330
[12002.517525]  ? get_signal+0x7f2/0x990
[12002.517528]  ? arch_do_signal_or_restart+0x64d/0x760
[12002.517531]  ? early_xen_iret_patch+0x5/0xc
[12002.517535]  ? exit_to_user_mode_prepare+0xd3/0x140
[12002.517538]  ? asm_exc_page_fault+0xc/0x30
[12002.517540]  ? irqentry_exit_to_user_mode+0x9/0x20
[12002.517542]  ? asm_exc_page_fault+0x22/0x30
[12002.517545]  ? early_xen_iret_patch+0x5/0xc
[12002.517547]  </TASK>
...
[12002.517625] CR2: 0000000000000344
[12002.517627] ---[ end trace 0000000000000000 ]---
[12002.517629] RIP: 0010:copy_fpstate_to_sigframe+0xad/0x330
[12002.517631] Code: 1f 44 00 00 48 8d bd 00 02 00 00 be 40 00 00 00 e8 b8 eb 58 00 48 85
c0 75 bc 65 48 8b 1c 25 c0 0b 02 00 65 81 05 9f 73 1e 5d <00> 02 00 00 48 8b 03 f6 c4 40
0f 85 ab 00 00 00 83 83 f8 1a 00 00
```

[12002.517632] RSP: 0000:ffffb298c2fb3df0 EFLAGS: 00050212

[12002.517634] RAX: 000000005d1e747a RBX: ffffb298c2fb3f58 RCX: 0000000000000008

[12002.517635] RDX: 0000000000000344 RSI: 0000000000000040 RDI: 00007fdb33e8ee80

[12002.517636] RBP: 00007fdb33e8ec80 R08: ffff9510262ef640 R09: 0000000000000000

[12002.517637] R10: 000000000000000b R11: 000000000000000a R12: ffff950c8fca5d40

[12002.517638] R13: ffff950c8fca4080 R14: ffff950c8fca4080 R15: 00007fdb33e8ec80

[12002.517639] FS: 00007fdb32000340(0000) GS:ffff95128f600000(0000) knlGS:0000000000000000

[12002.517641] CS: 0010 DS: 0000 ES: 0000 CR0: 0000000080050033

[12002.517642] CR2: 0000000000000344 CR3: 0000000204082000 CR4: 0000000000350ef0

Kernel Changelog:

commit a3c7c1a726a4c6b63b85e8c183f207543fd75e1b

Author: Juergen Gross <jgross@suse.com>

Date: Thu Jun 30 09:14:40 2022 +0200

    x86: Clear .brk area at early boot

    [Upstream commit 38fa5479b41376dc9d7f57e71c83514285a25ca0]

    The .brk section has the same properties as .bss: it is an alloc-only

    section and should be cleared before being used.

    Not doing so is especially a problem for Xen PV guests, as the

    hypervisor will validate page tables (check for writable page tables

    and hypervisor private bits) before accepting them to be used.

    Make sure .brk is initially zero by letting clear_bss() clear the brk

    area, too.

    Signed-off-by: Juergen Gross <jgross@suse.com>

    Signed-off-by: Borislav Petkov <bp@suse.de>

    Link: https://lore.kernel.org/r/20220630071441.28576-3-jgross@suse.com

    Signed-off-by: Sasha Levin <sashal@kernel.org>

And early_xen_iret_patch was added

https://github.com/torvalds/linux/commit/8b87d8cec1b31ea710568ae49ba5f5146
318da0d

Then later urgently removed:

> *x86/xen: Use clear_bss() for Xen PV guests*

https://github.com/torvalds/linux/commit/96e8fc5818686d4a1591bb6907e7fdb64
ef29884

And cleaned up in x86_urgent_for_v5.19_rc6:

https://github.com/torvalds/linux/commit/74a0032b8524ee2bd4443128c0bf97759
28680b0

CONFIG_XEN_VIRTIO was added:

https://github.com/torvalds/linux/commit/fa1f57421e0b1c57843902c89728f823a
bc32f02

Upstream: 38fa5479b41376dc9d7f57e71c83514285a25ca0

Fix: 36e2f161fb01795722f2ff1a24d95f08100333dd

Disclosure Timeline

- 2022-07-10 – Borislav Petkov & Juergen Gross fixes the vulnerability in
  mainline 5.19.rc6
- 2022-07-14 – Researcher encounters & reports vulnerability
  vulnerability on stable 5.18.11.

Links

https://sick.codes/sick-2022-128

https://github.com/sickcodes/security/blob/master/advisories/SICK-2022-
128.md

https://github.com/torvalds/linux/blob/babf0bb978e3c9fce6c4eba6b744c8754fd
43d8e/arch/x86/xen/mmu_pv.c#L1151

https://github.com/torvalds/linux/commit/8b87d8cec1b31ea710568ae49ba5f5146
318da0d

https://lore.kernel.org/all/20220308154317.815039833@infradead.org/

https://github.com/torvalds/linux/commit/96e8fc5818686d4a1591bb6907e7fdb64
ef29884

https://lore.kernel.org/all/20220630071441.28576-2-jgross@suse.com/

https://github.com/torvalds/linux/commit/74a0032b8524ee2bd4443128c0bf97759
28680b0

https://github.com/torvalds/linux/commit/fa1f57421e0b1c57843902c89728f823a
bc32f02

https://cdn.kernel.org/pub/linux/kernel/v4.x/ChangeLog-4.9.324

https://cdn.kernel.org/pub/linux/kernel/v4.x/ChangeLog-4.14.289

https://cdn.kernel.org/pub/linux/kernel/v4.x/ChangeLog-4.19.253

https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.4.207

https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.10.132

https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.15.56

https://cdn.kernel.org/pub/linux/kernel/v5.x/ChangeLog-5.18.13

Researchers

- Sick Codes https://github.com/sickcodes ||
  https://twitter.com/sickcodes
- Borislav Petkov, SUSE https://www.suse.com/
- Juergen Gross, SUSE https://www.suse.com/

CVE Links

https://sick.codes/sick-2022-128

https://github.com/sickcodes/security/blob/master/advisories/SICK-2022-128.md

https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2022-36123

https://nvd.nist.gov/view/vuln/detail?vulnId=CVE-2022-36123

Comments   11

Pingback: Vulnerabilidad en () - CVE-2022-36123 - Información y Soluciones

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - Adwareinfo.com

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - MalwareList.com

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - NoJoke.xyz

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - Malware

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - MalwareAlert.com

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - Adware.us

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS

users to cause a denial of service, or gain privileges. - WhatIsAdware.com

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - AdwareInformation.com

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - Spyware.ws

Pingback: CVE-2022-36123 - Linux kernel <5.18.13 lacks a certain clear operation for the block starting symbol (.bss). This allows Xen PV guest OS users to cause a denial of service, or gain privileges. - WhatIsSpyware.com

@sickcodes

@sickcodes

@sickcodes

Discord Server

sickcodes.slack.com

t.me/sickcodeschat

./contact_form