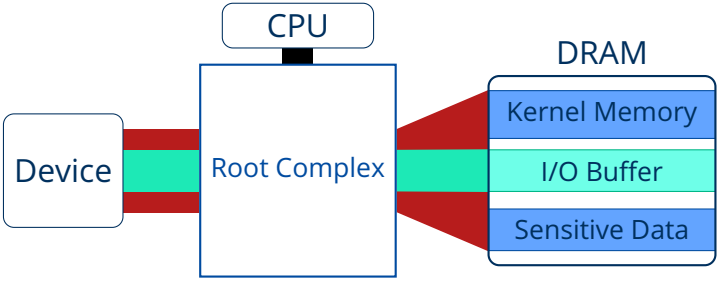
Christian Schwarz, Viktor Reusch, Maksym Planeta
Faculty of Computer Science, Institute of Systems Architecture, Chair for Operating Systems

DMA Security in the Presence of IOMMUs

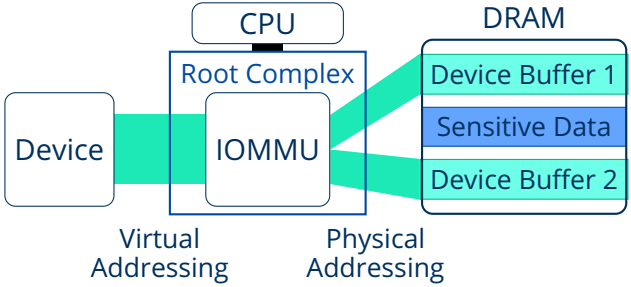
Presentation at the *Frühjahrstreffen der Fachgruppe Betriebssysteme* 2022 // Hamburg, March 18, 2022

Unprotected DMA



- All PCIe Devices have full main memory access via DMA
- Faulty or malicious devices can harm the system

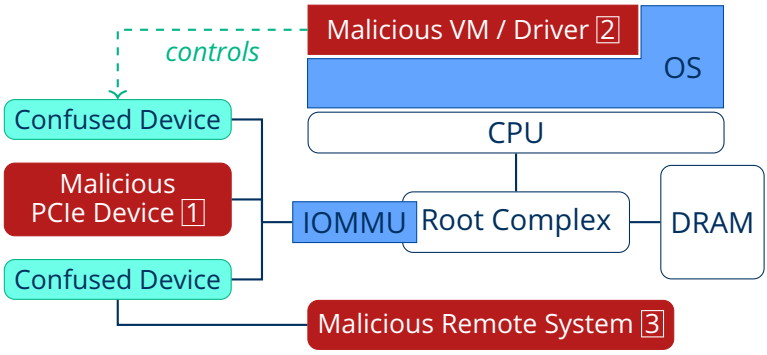
IOMMU Protection



- The IOMMU puts PCIe devices into address spaces
- Numerous benefits, similar to virtual memory for userspace processes
- Unfortunately, it does not solve all security issues

Table of Contents

Attack Vectors



Malicious Devices

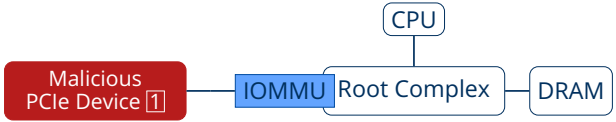
Malicious VM

Malicious Remote System

Conclusion

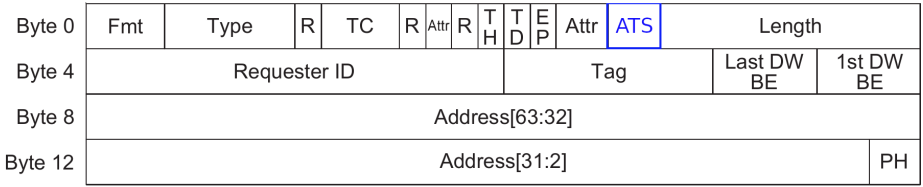
Thunderbolt (Ruytenberg: Thunderspy)

[1]



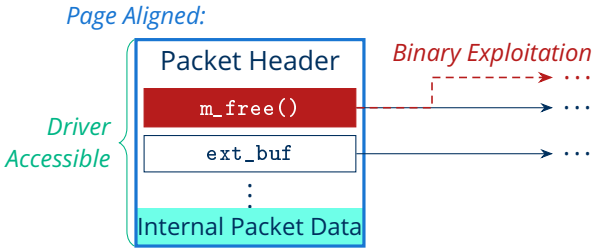
- Thunderbolt bypasses the IOMMU without the "Kernel DMA Protection" BIOS Setting (exists since 2019)
- Apple hardware is not affected

ATS (Markettos et al.: Thunderclap)



- Address Translation Services (ATS) allow disabling the IOMMU (Linux disabled this for "untrusted devices" after 4.20), also: `pci=noats`
- Thunderbolt bypasses the IOMMU without the "Kernel DMA Protection" BIOS Setting (exists since 2019)

Subpage Granularity (Markettos et al.: Thunderclap)



- The IOMMU always guarants whole pages
- Evil Device can impersonate any existing device → attack any driver
- Thunderclap: ACE on FreeBSD, and MacOS (below 10.12.4)
- Linux "only" leaked data (`sk_buf` used `kmalloc`) → SPADE, KASAN

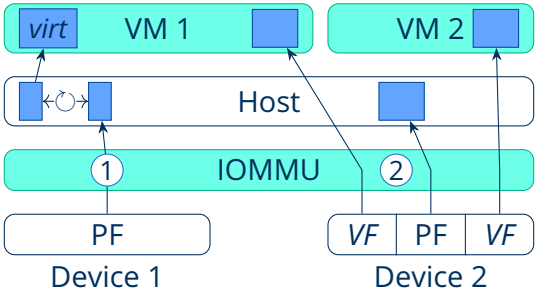
Malicious Devices

Malicious VM

Malicious Remote System

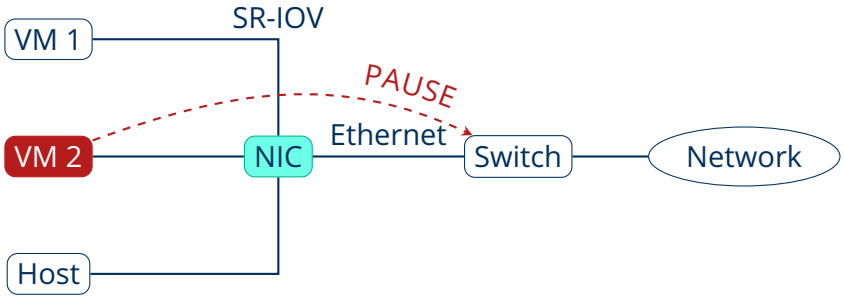
Conclusion

VM Device Sharing using the IOMMU



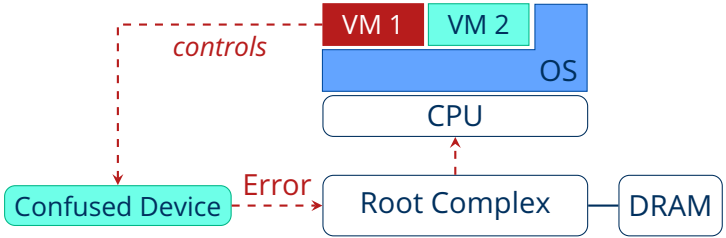
1. Emulate a virtual device in software
2. Map I/O Buffers for a *virtual function* ($\rightarrow$ *SR-IOV*) directly to the VM

Ethernet FC/OAM (Smolyar et al.: Securing Self-Virtualizing Ethernet Devices)



- Ethernet Flow Control (FC): the VF can send a `PAUSE` request
- Ethernet operations, administration and maintenance (Ethernet OAM): the VF can send a `link fault` → link gets disconnected
- Affected hardware can't safely use SR-IOV

PCIe Error Handling (Khattri et al.: PCIe Device Attacks: Beyond DMA)



- PCIe devices can emit "Uncorrectable Fatal Errors" to the Root Complex
- The OS might decide to reset the whole link
- If a VF can trigger these → DOS
- E.g. Intel 700 Series Ethernet Controllers → CVE-2019-0144

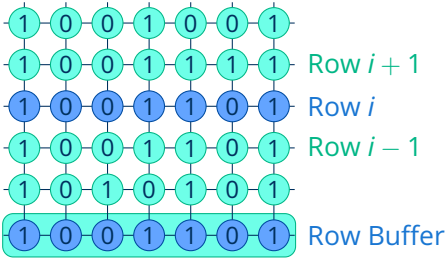
Malicious Devices

Malicious VM

Malicious Remote System

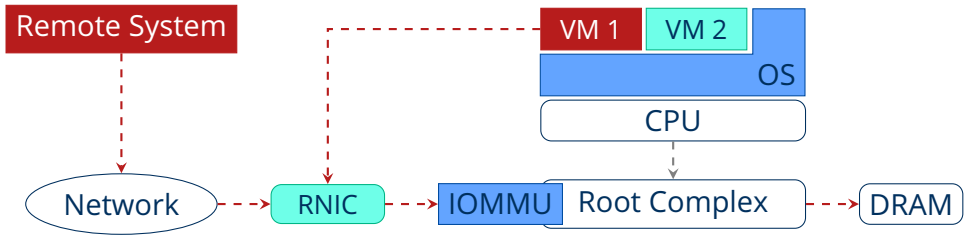
Conclusion

Rowhammer (Tatar et al.: Throwhammer)



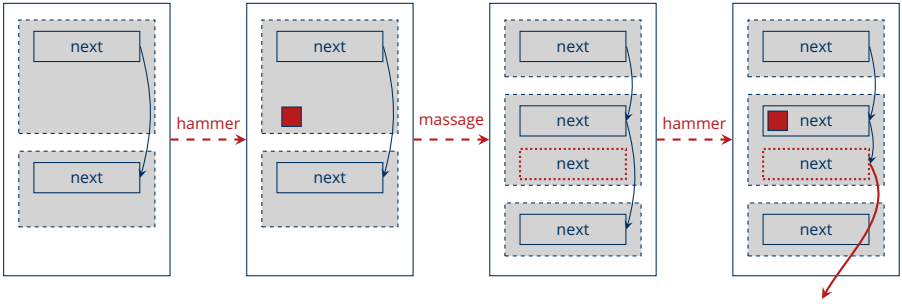
- Read $r-1$ and $r+1$ repeatedly $\rightarrow$ bitflip in i
- DDR4 Protections can still be broken, DDR5 seems to fare better
- Software mitigations (like ANVIL) don't detect DMA

RDMA Networking



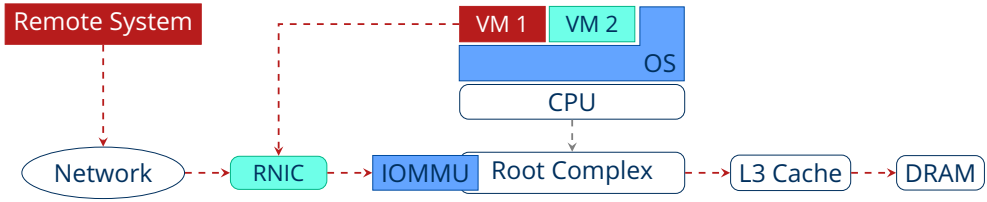
- High bandwidth, low latency, kernel bypass (once established)
- Multiple implementations (RoCE, RoCE v2, iWARP, Infiniband, …)

Rowhammer through RDMA (Tatar et al.: Throwhammer)



1. Read memory and observe where bitflips occured
2. *Massage* memory to move critical pointers into the bitflipped locations
3. Send the same RDMA requests again and hope for the same flips
4. Trigger access to the corrupted locations ⇝ ACE e.g. on RDMA memcached

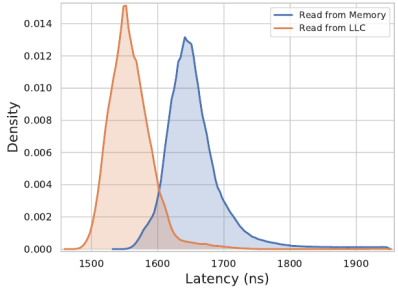
Cache Side-Channel Attacks (Kurth et al.: NetCAT)



- Data Direct IO: applicable for Xeon E5 and E7 v2 processor families
- Parts of the L3 cache (LLC) are used for DMA

Cache Side-Channel Attacks (Kurth et al.: NetCAT)

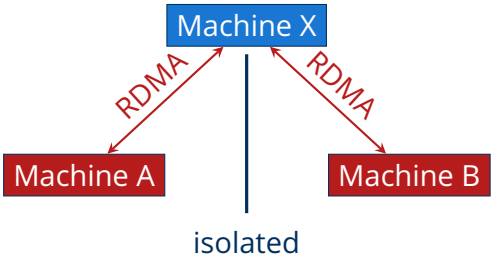
Observable timing differences for RDMA reads:

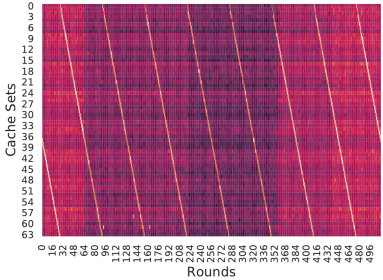


Measured on a Xeon Silver 4411 using ConnectX-4 Infiniband [3]

Even over RDMA, the faster access times for cached data are observable

Cache Side-Channel Exploits (Kurth et al.: NetCAT)





[3]

- Covert channel between isolated machines by observing evictions
  → 145 KB/s
- Observing network packet times by locating receive ring buffers
  → e.g. SSH keystroke prediction

Conclusion

- The PCIe Bus should be treated like an untrusted public network.
- DMA Security is a known concern, devices and software are improving.

- Disable Thunderbolt on older hardware
  (check for Kernel DMA Protection)
- Replace SR-IOV for the networking of untrusted VMs
  (unless DOS is tolerable or Ethernet FC and OAM are filtered out)
- Disallow RDMA for all untrusted sources
  (unless ECC DDR4 or DDR5 RAM is used and there is no DDIO)

Christian Schwarz, Viktor Reusch, Maksym Planeta
Faculty of Computer Science, Institute of Systems Architecture, Chair for Operating Systems

DMA Security in the Presence of IOMMUs

Presentation at the *Frühjahrstreffen der Fachgruppe Betriebssysteme* 2022 // Hamburg, March 18, 2022

References I

[1] Amin. *Thunderbolt 3 interface USB-C ports*. via Wikimedia Commons, CC BY-SA 4.0. 2018. URL: https://commons.wikimedia.org/wiki/File:Thunderbolt_3_interface_USB-C_ports.jpg.

[2] Hareesh Khattri, Nagaraju N Kodalapura (Raju), and Nam N Nguyen. *PCIe Device Attacks: Beyond DMA*. Intel Coorperation. 2021. URL: https://i.blackhat.com/USA21/Wednesday-Handouts/us-21-PCIe-Device-Attacks-Beyond-DMA-Exploiting-PCIe-Switches-Messages-And-Errors.pdf (visited on 11/22/2021).

[3] Michael Kurth et al. "NetCAT: Practical Cache Attacks from the Network". In: *2020 IEEE Symposium on Security and Privacy (SP)*. 2020, pp. 20–38. DOI: 10.1109/SP40000.2020.00082.

References II

[4] A. Markettos et al. “Thunderclap: Exploring Vulnerabilities in Operating System IOMMU Protection via DMA from Untrustworthy Peripherals”. In: Jan. 2019. DOI: 10.14722/ndss.2019.23194. URL: https://thunderclap.io/thunderclap-paper-ndss2019.pdf.

[5] Björn Ruytenberg. *Breaking Thunderbolt Protocol Security: Vulnerability Report*. 2020. URL: https://thunderspy.io/assets/reports/breaking-thunderbolt-security-bjorn-ruytenberg-20200417.pdf.

[6] Igor Smolyar, Muli Ben-Yehuda, and Dan Tsafrir. “Securing Self-Virtualizing Ethernet Devices”. In: *Proceedings of the 24th USENIX Conference on Security Symposium*. SEC’15. Washington, D.C.: USENIX Association, 2015, pp. 335–350. ISBN: 9781931971232.

References III

[7] Andrei Tatar et al. “Throwhammer: Rowhammer Attacks over the Network and Defenses”. In: *Proceedings of the 2018 USENIX Conference on Usenix Annual Technical Conference*. USENIX ATC ’18. Boston, MA, USA: USENIX Association, 2018, pp. 213–225. ISBN: 9781931971447.