Virtualization Based Hardening: Securing Container Workloads and beyond

Andrei LUTAS

Senior Security Researcher, Bitdefender

Jun NAKAJIMA

Senior Principal Engineer, Intel Corporation

October 2019 – KVM Forum – Lyon, France

About the speakers

- *Andrei LUTAS*
  - *HVI lead developer*
  - *>11 years experience*
  - *Low-level / security enthusiast*
- *Jun NAKAJIMA*
  - *Virtualization Architect*
  - *> 15 years experience*

Intel Contributors: ***Rong Liu and Sainath Grandhi***

Bitdefender

intel

Agenda

- *About VBH*
- *About HVI*
- *VBH + HVI Architecture*
- *CVE-2016-5195*
- *CVE-2017-7308*
- *CVE-2019-5736*
- *Conclusions*

Bitdefender

intel

About VBH

*H/W Virtualization features allow **Ring -1**, or **Hypervisor**:*

- *Virtualize physical memory, CPU (control registers, MSRs, etc.), I/O,*
- *Monitor, isolate, and protect resources*

*VBH utilizes them to protect **bare metal kernel**:*

- *Initially presented at KVM Forum 2016**
- *Thin hypervisor loaded in the Linux kernel*
- *De-privileges bare-metal Linux*
- *Provides additional protection to the kernel without visible overhead*

Bitdefender

*: http://www.linux-kvm.org/images/4/40/01x05-Jun_Nakajima-Kernel_Protection_Using_Hardware-Based_Virtualization.pdf

intel

About HVI

"The approach of inspecting a VM from the outside for the purpose of analyzing the software running inside it"

Garfinkel & Rosenblum, 2003

Bitdefender

intel

About HVI

- *HVI == Hypervisor Introspection*
- *Usually runs **outside** the protected operating system*
- ***Bridges** the semantic gap*
- *Leverages virtualization features to **provide protection***
- *In a regular scenario – protects the **kernel** and the **applications** against attacks*
- *HVI for **containers** aims to protect the host kernel against malicious containers*

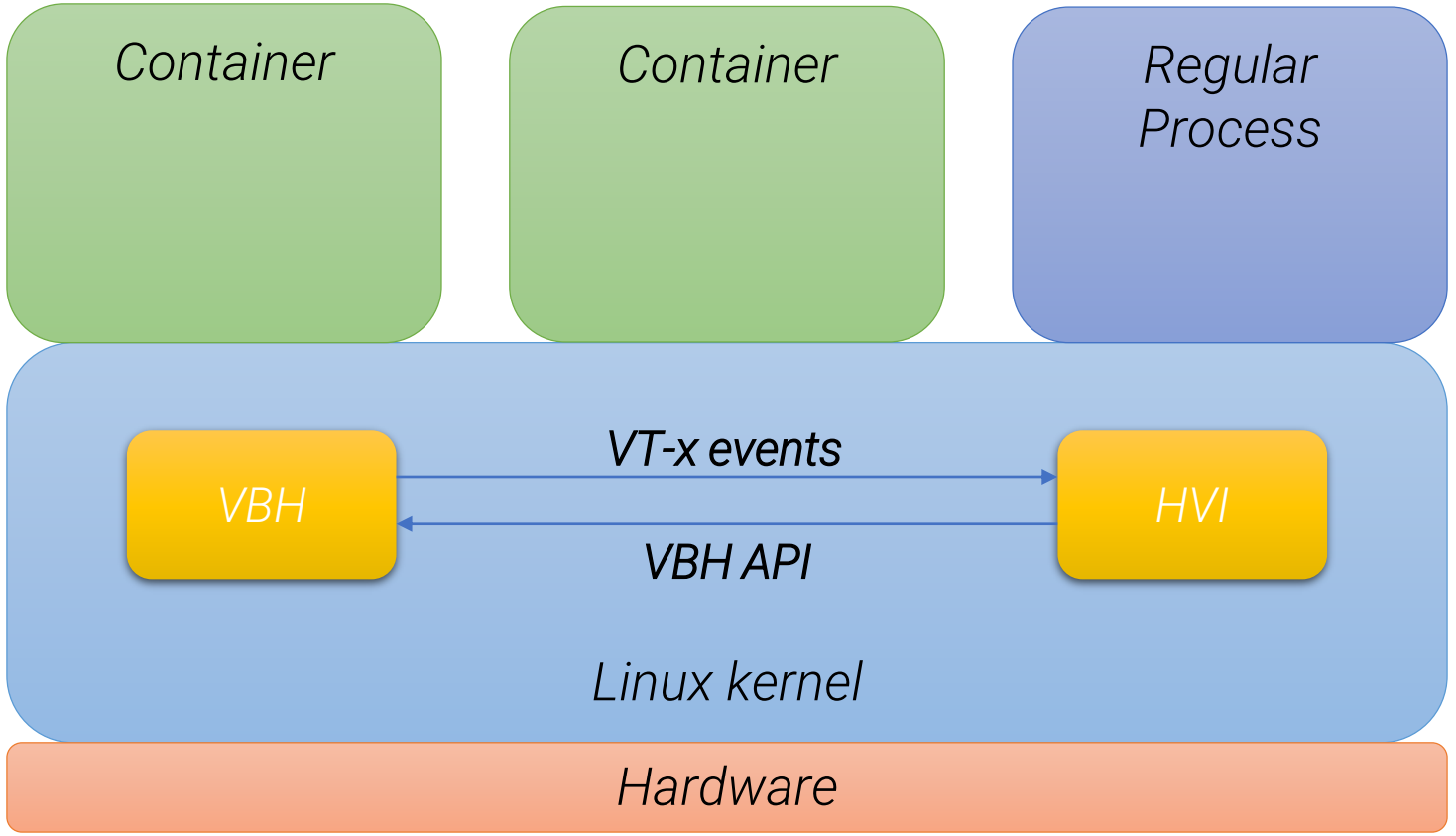
Bitdefender

intel

About HVI

- *Leverage EPT (Extended Page Tables) to provide memory protection*
  - *Protect stacks, heaps, etc. against **execution** (exploit)*
  - *Protect code, read only pages, critical structures against **writes***
- *Leverage other VT-x features, in order to:*
  - *Protect critical **MSRs** against malicious writes (ie, **SYSCALL**)*
  - *Protect **CRs** against malicious modifications (ie, **CR4.SMEP**, **SMAP**)*
  - *Protect **descriptor tables** against modifications*

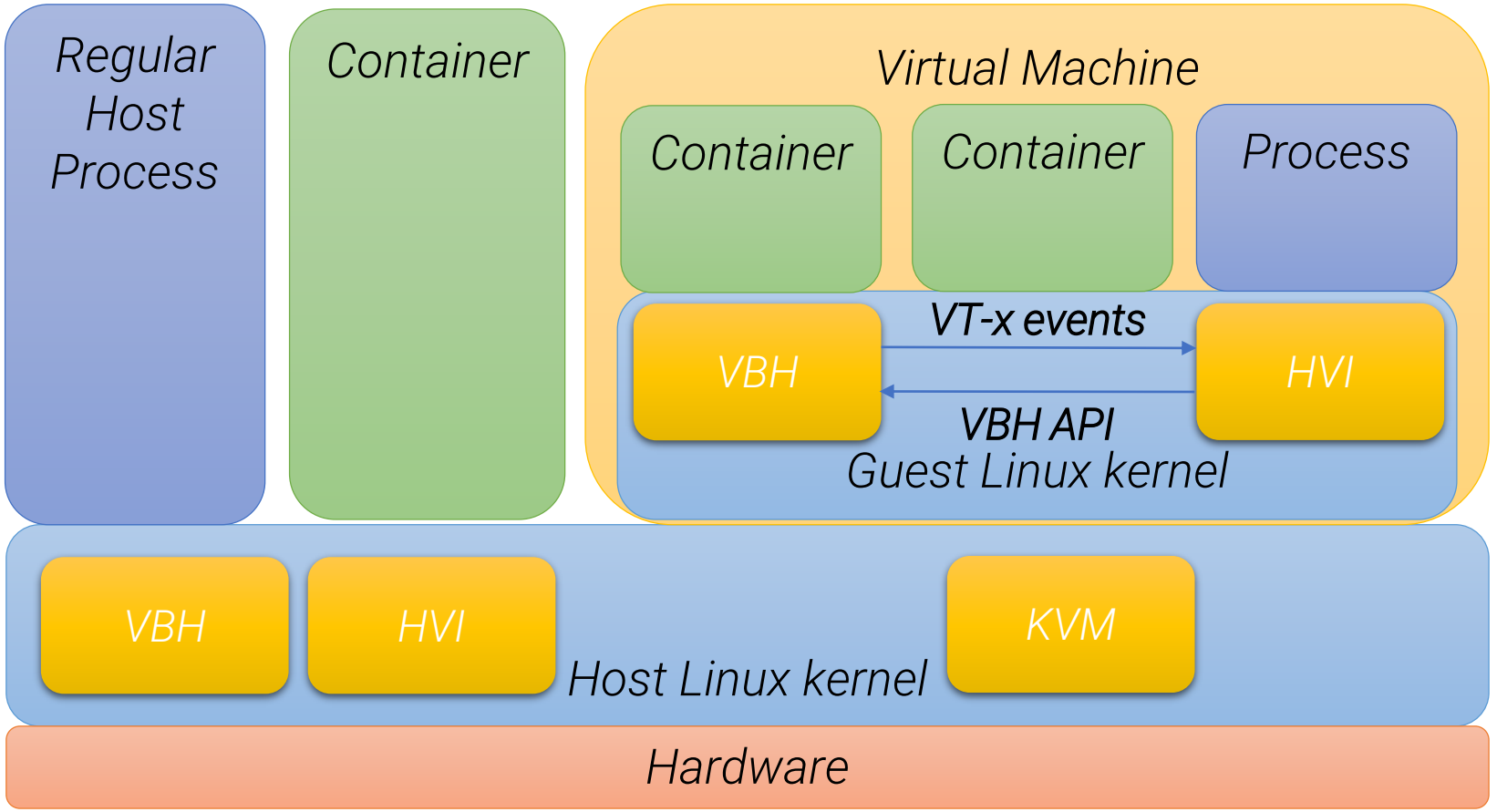
VBH + HVI Architecture – bare-metal



Bitdefender

intel

VBH + HVI Architecture - nested



Bitdefender

intel

CVE-2016-5195 (aka DirtyCOW)

- *A race condition in mm/gup.c in the Linux kernel 2.x through 4.8.3, allows local users to gain privileges by leveraging the incorrect handling of a copy-on-write (COW) feature to write to a read-only memory mapping. This has been exploited in the wild in October 2016 (Dirty COW).*
- *Remove EPT write permissions for the vDSO page*

CVE-2016-5195 (aka DirtyCOW)

- *Demo*

CVE-2017-7308

- *It was found that the packet_set_ring() function of the Linux kernel's networking implementation did not properly validate certain block-size data. A local attacker with CAP_NET_RAW capability could use this flaw to trigger a buffer overflow resulting in a system crash or a privilege escalation.*
- *Prevent CR4.SMEP & CR4.SMAP from being cleared*

CVE-2017-7308

- *Demo*

Bitdefender

intel

CVE-2019-5736 (aka runc)

- *runc through 1.0-rc6, as used in Docker before 18.09.2, and also in other products, allows attackers to overwrite the host runc binary and consequently obtain host root access.*
- ***Hook the file open function and prevent writes inside runc***

CVE-2019-5736 (aka runc)

- *Demo*

Bitdefender

intel

Conclusions

- *We demonstrated how several container attacks have been successfully blocked by VBH + HVI*
  - *Beyond security – debug/monitoring tool*
- *Intel ® and Bitdefender joined forces in an open-source effort to secure containers*
  - *https://github.com/intel/vbh*
  - *https://github.com/bitdefender/vbh_sample*
- *We want to create an ecosystem around VBH & container security*
  - *The community can contribute to both VBH and HVI*

Bitdefender

intel

Questions

Bitdefender

intel