KVM Status Report

KVM Forum 2021

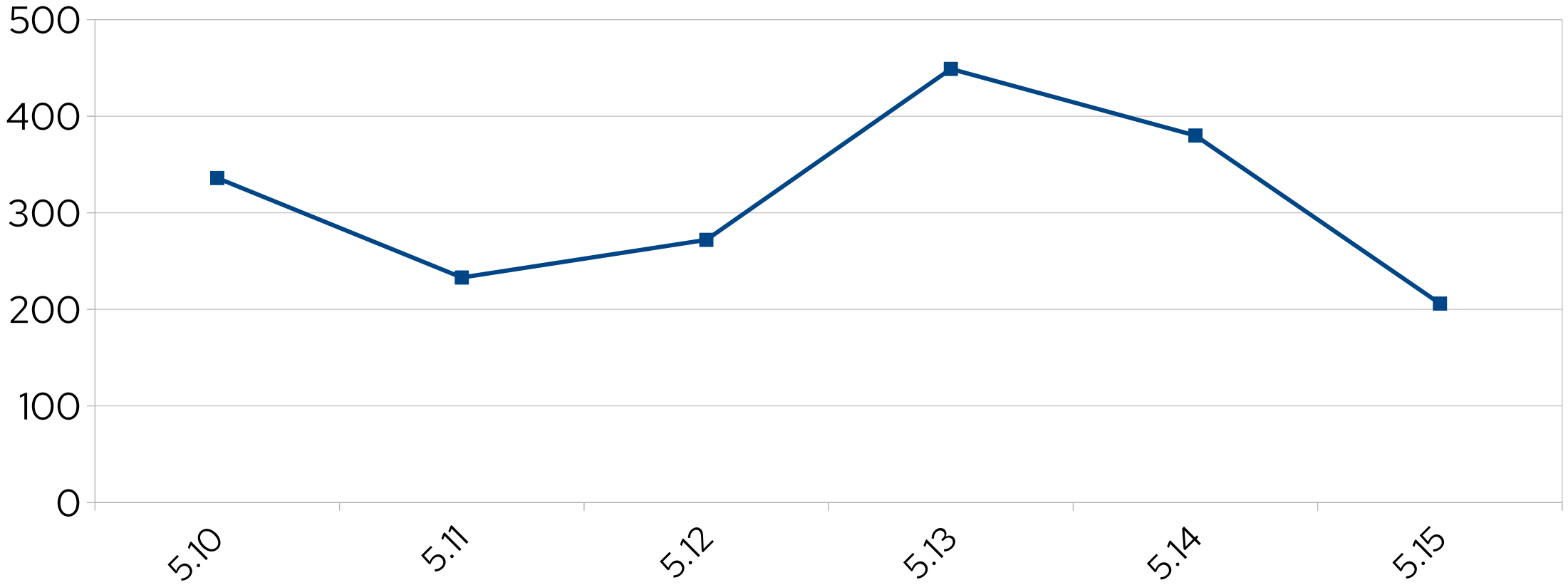
Paolo Bonzini, Red Hat
Distinguished Engineer

Red Hat

October 2020 – September 2021

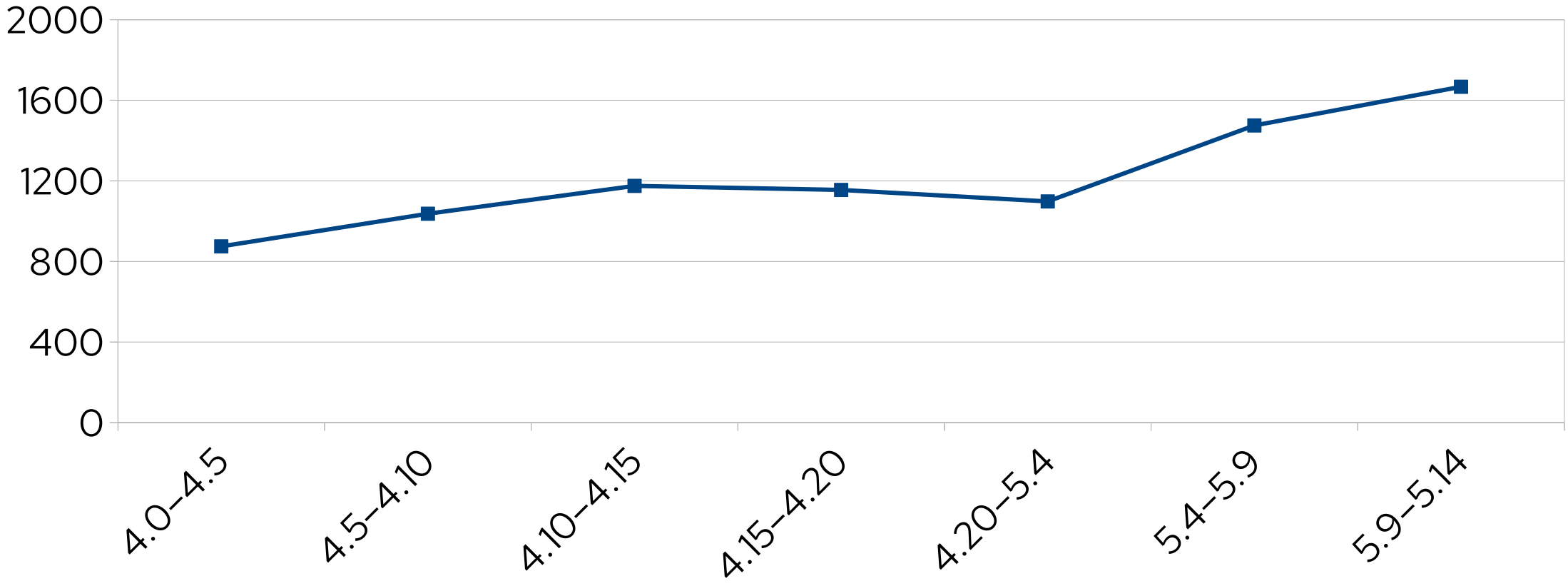
- October 2020: Linux 5.9
- September 2021: Linux 5.15-rc1
- 5 releases, one more under way
- 1911 commits (5.9-rc1 to 5.15-rc1)
- 168 commits destined to stable releases

Commits in each release



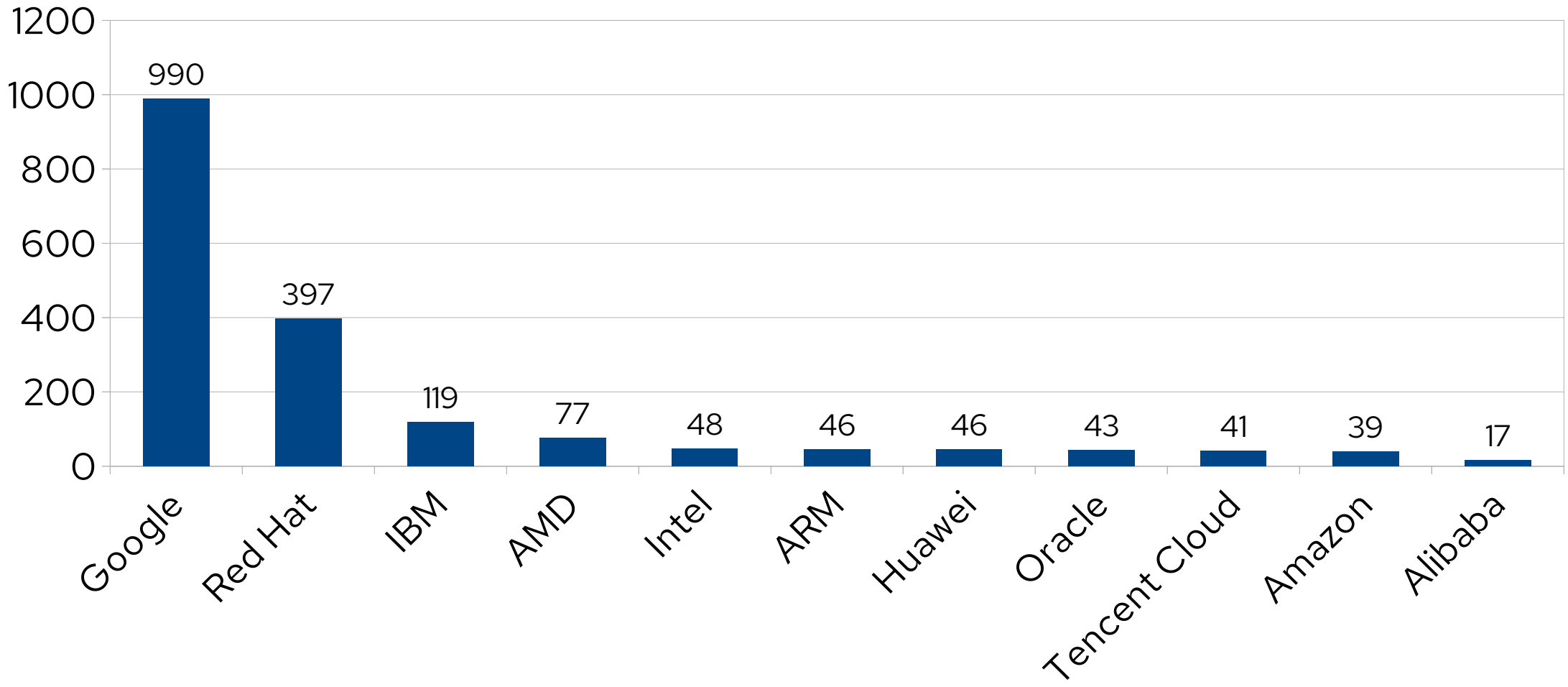
Red Hat

Commits in each group of 5 releases



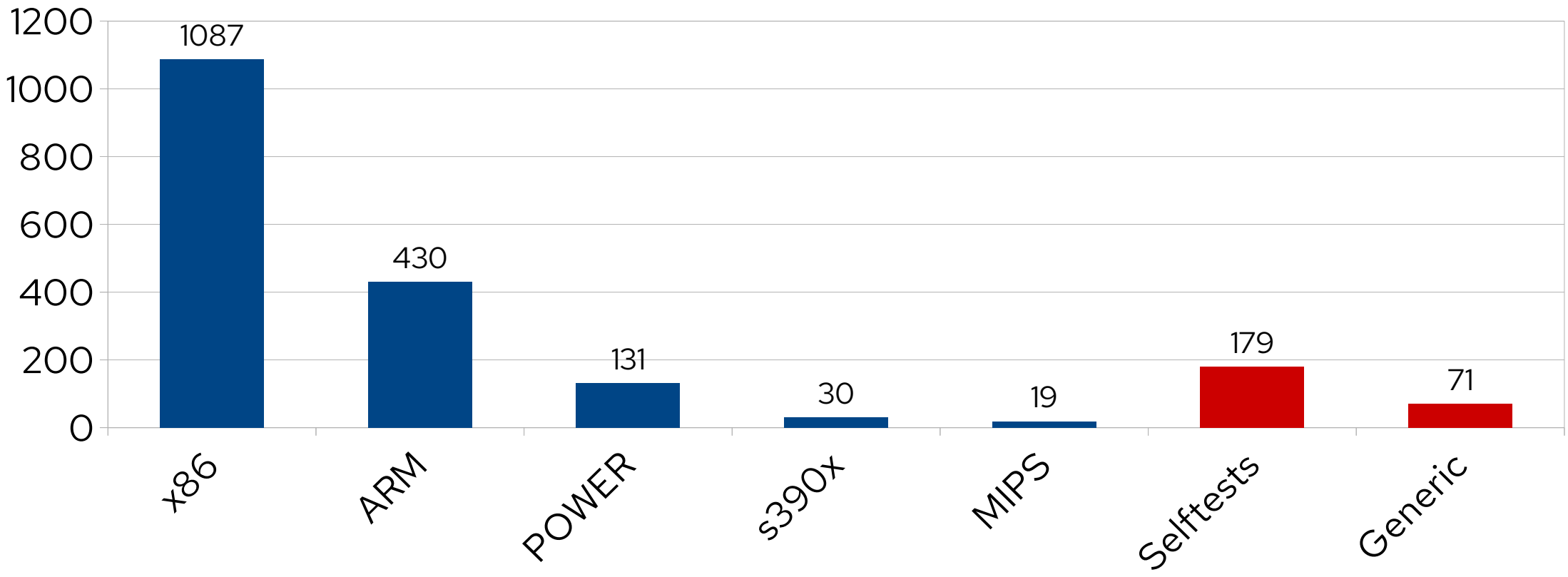
Red Hat

Commits by employer since 5.9-rc1



Red Hat

Commits by architecture since 5.9-rc1



Red Hat

Highlights: all architectures

- MMU notifier optimizations (5.13, 5.15)
- Binary statistics API (5.14)
- Selftests: +10.000 lines!

x86: MMU

- New MMU (5.10–5.15, KVM Forum 2019)
  - Parallel page faults
  - Lazy rmap allocation
  - Enabled by default in 5.15
- Dirty page ring buffer (5.11, KVM Forum 2020)

x86: Confidential VMs

- SEV-ES support (guest 5.10, host 5.11)
- SEV shared encryption context (5.13)
- SEV live migration (5.13, KVM Forum 2019)
- Work on TDX under way (KVM Forum 2020)

x86: Nested virtualization

- Live migration stability (Intel, AMD)
- Wait-for-SIPI activity state (Intel)
- Nested TSC scaling (Intel)
- Optimizations for Hyper-V hosts (AMD)
- Security fixes (AMD)
- Shoutout to GSoC “SVM emulation on TCG” project

x86: Miscellaneous

- MSR exiting to userspace (5.10)
- Xen interface emulation (5.12)
- Static calls (5.12, KVM Forum 2019)
- Improvements to yield-on-spin heuristics (5.13)
- Selective enabling of Hyper-V hypercalls (5.14)

ARM: Protected KVM

- Presented at KVM Forum 2020 (see also LWN coverage)
- Linux (EL1) is isolated from KVM (EL2)
  - Guests not visible to EL1
  - Isolate various “trusted” binary blobs

ARM: Protected KVM

- Host kernel stage-2 protection (5.13)
- Hypervisor isolation (5.13)
- Hypervisor tracking of page ownership (5.15)
- Working on guest-host page sharing for device DMA
- Expecting to ship in Android 13!

What's next?

- TDX, SEV-SNP
- More nested virtualization!
  - AMD nested TSC scaling
  - AMD "nested nested" (vGIF, vVMLOAD/VMSAVE)
- More MMU
  - Eager page tables for SEV-SNP and TDX
  - Remove legacy two-dimensional paging?
- Binary statistics for QEMU and libvirt

Thank you

linkedin.com/company/red-hat

youtube.com/user/RedHatVideos

facebook.com/redhatinc

twitter.com/RedHat

Red Hat