

KVM Status Report

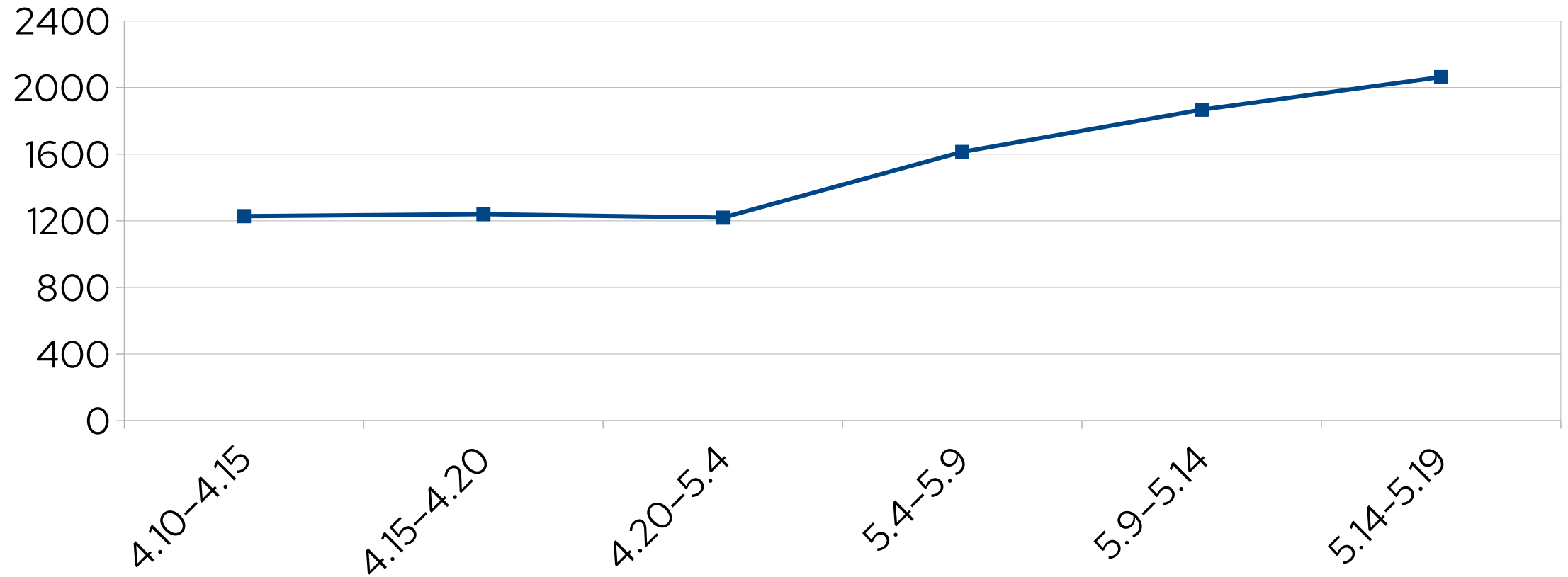
KVM Forum 2022

Paolo Bonzini, Red Hat
Distinguished Engineer

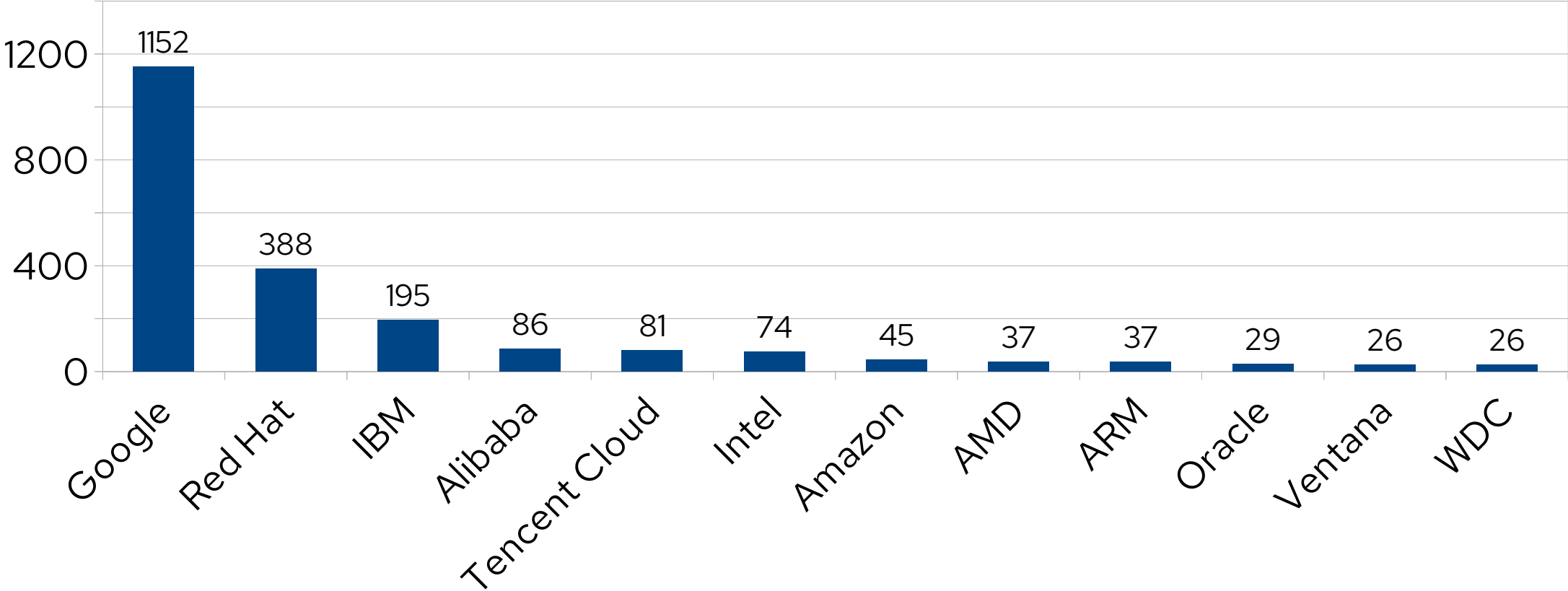
September 2021 – September 2022

- September 2021: Linux 5.15-rc1
- September 2022: Linux 6.0 well under way
- (Almost) 6 releases
- ~2500 commits; 202 to stable releases
 - Up from 2200 and 180 between 5.9-rc1 and 5.15-rc1
 - Ratio roughly unchanged, a little less than 1:12
- Welcome RISC-V to upstream KVM! (5.16)

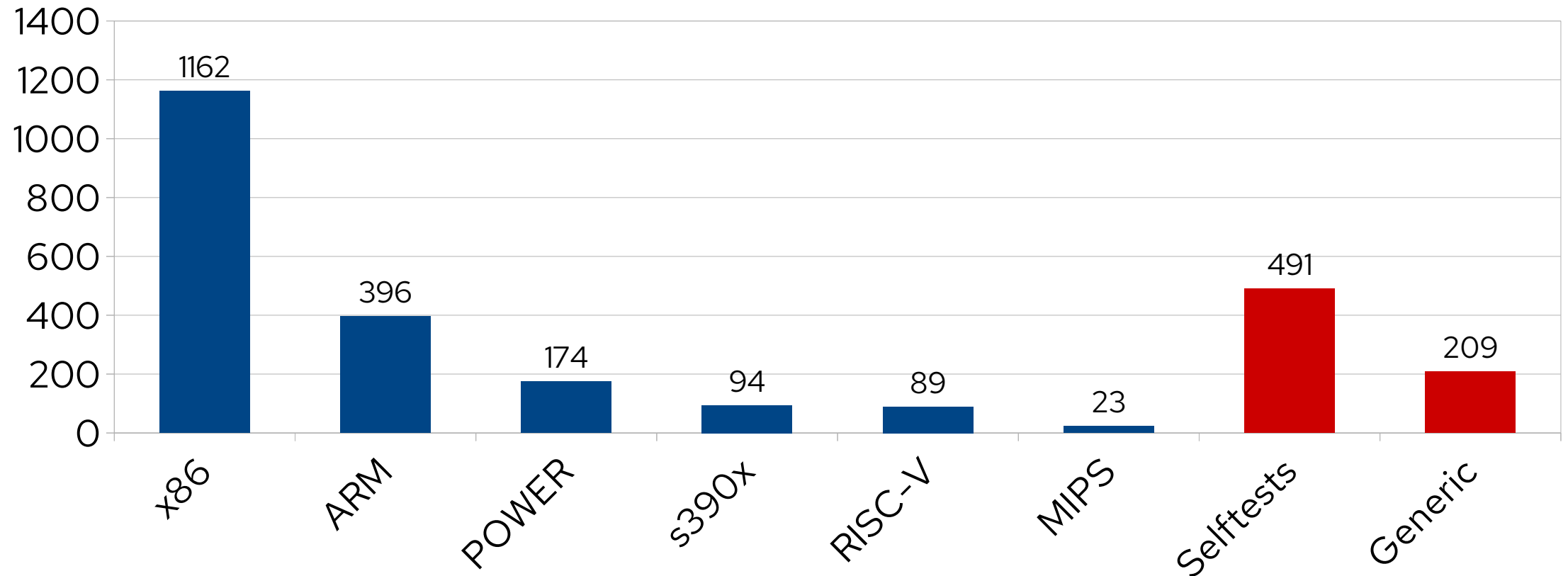
Commits in each group of 5 releases



Commits by employer since 5.15-rc1



Commits by architecture since 5.15-rc1



Use Linux library code more

- Memslot lookup (interval tree)
- vCPU lookup (xarray)
- x86 page table destruction (workqueue)

x86 highlights

- API for continuous TSC over migration
- Many APICv/AVIC cleanups
- Many more MMU cleanups and fixes
- In-kernel Xen event channel delivery
- Eager splitting of page tables

x86 hardware features - AMD

- Nested LBR
- Nested TSC scaling
- “Nested nested” acceleration: vGIF, vVMLOAD/VMSAVE
- AVIC with physical APIC ID > 255 (aka “x2AVIC”)

x86 hardware features - Intel

- AMX and dynamic XSAVE states, thanks Thomas Gleixner!
- Intel IPI virtualization
- Intel PEBS (Precise Event-Based Sampling) virtualization

Arm highlights

- Support for timed event wait instructions WFxT
- Support for asymmetric PMU setup
- Apple M1 support
- PSCI-based suspend
- Hypercall selection from userspace
- New VMID allocator using fewer IPIs
- Hypervisor stack guard pages and stack traces

Arm: Protected KVM

- EL2 filtering of system registers
- Limitation of some hypercalls after initialization
- Selective sharing of pages from EL1 to EL2

s390

- Secure guests
 - Lazy destroy of secure VMs
 - Ultravisor communication device driver
 - Adapter interrupt virtualization
- Storage key improvements/fixes

Maintenance changes

- POWER changes go through architecture tree
- New x86 maintainers sending pull requests to me
 - Sean Christopherson
 - Vitaly Kuznetsov

What's next? x86

- Improved CI
- x86 confidential computing
 - Intel TDX
 - AMD SEV-SNP

What's next? Feature parity

- Paravirtualization features
 - Asynchronous page faults
 - Steal time
 - Proposals worked on in the RISC-V hypervisor SIG
- Hypervisor features
 - Dirty page ring
 - Scalable MMU

Thank you



[linkedin.com/company/red-hat](https://www.linkedin.com/company/red-hat)



[youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)



[facebook.com/redhatinc](https://www.facebook.com/redhatinc)



twitter.com/RedHat