

NEMU

The No EMUlation Hypervisor

September 11th, 2018

NEMU

- **No EMUlation**
- QEMU based
- Exclusively focused on cloud-specific workloads
- Removes all features, platform and hardware emulation not required for the cloud
- Keeps the performance, stability and robustness of QEMU

NEMU Goals

- Small attack surface (Low complexity, low footprint)
- High performance (Comparable to QEMU)
- No hardware emulation
- UEFI firmware only
- Linux and modern Windows guests only
- x86-64 and aarch64 only
- *Hardware-Reduced* ACPI
- Hotplug: CPU, memory, NVDIMM and PCI

Initial Proof Of Concept (May 2018)

- pc and q35 (i386), virt (ARM) machine types only
- Semi-manual code removal: features, platforms and virtual devices
- 75% code size reduction
- Binary size reduced from 12.5 MB to 4.3 MB
- Shared libraries reduced from 97 to 29
- Device model reduced from 236 to 66 devices
- Published on GitHub as “[experiment/code-reduction](#)”

NEMU Today (September 2018)

- New machine type on top of QEMU 3.0.0: ***i386/virt***
- *Hardware-Reduced* ACPI
- UEFI firmware only (OVMF)
- Minimized hardware emulation
- No legacy hardware support
- Minimized device model
- ACPI-based hotplug
- Integrated and extended CI
- Published as “[topic/virt-x86](#)”

Hardware-Reduced ACPI

- *i386/virt* complies with the *Hardware-Reduced ACPI* specification
- Specifically designed for modern, legacy-free and UEFI-based platforms
- Significantly less complex ACPI core code
- Compatible with *aarch64* kernels
- Needs a [kernel patch](#) to support hotplug

Firmware

- One single virtual UEFI firmware: OVMF
- *i386/virt* support added to OVMF
- OVMF changes:
 - Replace fixed function ACPI timer with KVM clock + TSC
 - Don't use CMOS to get memory details
 - Use KVM clock instead of emulated RTC
- Temporary [OVMF fork](#)

Minimized Hardware Emulation

- Minimal PCI host bridge emulation (pci-lite)
- No chipset-specific emulation (LPC, PCH, MCH)
- No ISA, SMBUS or RTC/CMOS emulation
- Clock and IRQ controller offloaded to KVM
- Virtual ACPI device for hotplug/shutdown/reset support

Hotplug

- CPU, memory, NVDIMM and PCI devices
- Not a common use case for cloud workloads
- Mostly needed for VM based containers support (Kata Containers)
- Purely ACPI based (even for PCI devices)

Continuous Integration

- NEMU Automatic Test System (NATS)
- Programmatic approach to QEMU functional testing
- Jenkins-based CI, each GitHub PR is tested

Status

- NEMU is already an open source project
- Temporary OVMF fork to support the new ***i386/virt*** machine type
- The ***i386/virt*** machine type boots UEFI-based Linux cloud workloads (Clear Linux, Ubuntu Xenial & Bionic)
- The ***i386/virt*** machine type runs Kata Containers

Plans

- Upstreaming of *virt* machine type and OVMF support
- Documentation
- Refactoring of build system to allow more minimal platform configuration
- Automated (and manual) code removal/reduction of everything not part of “virt platform”
- CI extensions (More features tested, ARM support)
- Work toward legacy-free PCI
- Kata Containers integration

Questions ?