Performance Profiling in a Virtualized Environment

Jiaqing Du[+], *Nipun Sehrawat*[*], *Willy Zwaenepoel*[+]

[+]EPFL, Switzerland

[*]IIT Guwahati, India

Virtualization-based Clouds

- Diverse implementations
  - software: Xen, KVM, VMware, …
  - hardware: Intel x86, AMD x86, PowerPC, …
- Opportunities for performance profiling & tuning
  - public cloud: guest OS & applications
  - private cloud: whole software stack

Profilers based on CPU performance counters do not work well with virtual machines.

Profiling in a Virtualized Environment

- General challenge
- XenOProf: profiling in the VMM
  - only for paravirtualization-based Xen
  - require accesses to the VMM
- Profiling in the guest
  - normally no results
  - *"OProfile can't work with VMware when using performance counter interface."*

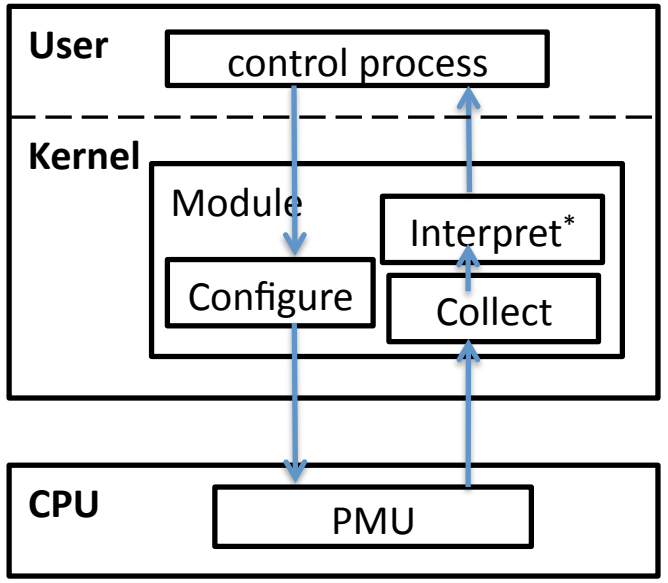
Performance Profiling

- Understand runtime behavior
- Tune performance
- Mature & used extensively
  - VTune, OProfile, …

%CYCLE	Function	Module
98.5529	vmx_vcpu_run	kvm-intel.ko
0.2226	(no symbols)	libc.so
0.1034	hpet_cpuhp_notify	vmlinux
0.1034	native_patch	vmlinux

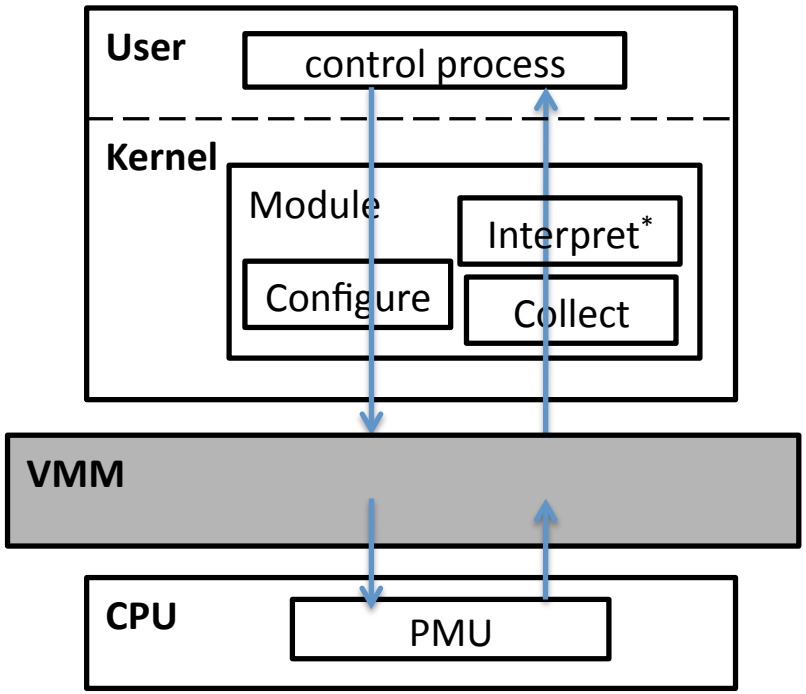
Native Profiling

- Performance monitoring unit (PMU)
  - a set of event counters
  - generate an interrupt when a counter overflows
- PMU-based profiler
  - sampling configuration
  - sample collection
  - sample interpretation



Guest-wide Profiling

- Expose PMU interfaces to the guest



Guest-wide Profiling

- Where to save & restore the registers?
- CPU switch
  - only in-guest execution is accounted to the guest

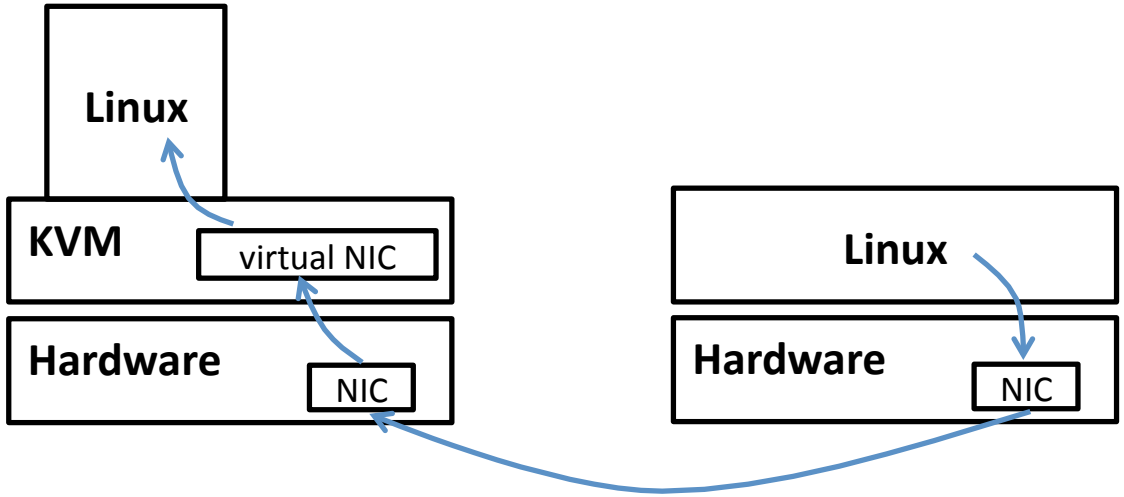
- Domain switch
  - in-VMM execution is also accounted to the guest

Guest-wide Profiling for KVM

- Kernel-based virtual machine (KVM)
  - a Linux kernel subsystem
  - a set of kernel modules + QEMU
  - built on hardware virtualization extensions
- Intel VT extensions
  - provide a list of hardware features
  - facilitate our implementation
- No modifications to the guest and the profiler

Profiling Packet Receive

- Experiment
  - push packets to a Linux guest in KVM
  - run OProfile in the guest
  - monitor instruction retirements



Profiling Packet Receive

CPU Switch

INSTR	Function
167	csum_partial
106	csum_partial_copy_generic
74	copy_to_user
47	ipt_do_table
38	tcp_v4_rcv
…	…
19	cp_rx_poll
6	cp_start_xmit
6	cp_interrupt
3	native_apic_mem_write

1184 counter overflows

Domain Switch

INSTR	Function
2261	cp_interrupt
1336	cp_rx_poll
1034	cp_start_xmit
421	native_apic_mem_write
374	native_apic_mem_read
191	csum_partial
105	csum_partial_copy_generic
94	copy_to_user
79	ipt_do_table
51	tcp_v4_rcv

7286 counter overflows

Other Things in the Paper

- System-wide profiling
  - profiling in the VMM
  - provide full-scale view: guest + VMM
- Virtualization techniques
  - paravirtualization
  - dynamic binary translation

Conclusions

- Profilers do not work well with virtual machines.
- We implement guest-wide profiling in a VMM based on hardware assistance.
- Profiling helps understand the real cost of I/O operations in a guest.