

COMPUTER SCIENCE RESEARCH SEMINAR

Overcoming Virtualization Overheads for Large-vCPU Virtual Machines

Ozgur Kilic, PhD Student

Department of Computer Science, Binghamton University

Friday, February 22, 2019 at noon in room R15, Engineering Building

Abstract: Virtual Machines (VM) frequently run parallel applications in cloud environments, and high-performance computing platforms. It is well known that configuring a VM with too many virtual processors (vCPUs) worsens application performance due to scheduling cross-talk between the hypervisor and the guest OS. Specifically, when the number of vCPUs assigned to a VM exceeds available physical CPUs then parallel applications in the VM experience worse performance, even when number of application threads remains fixed. In this paper, we first track the root cause of this performance loss to inefficient hypervisor-level emulation of inter-vCPU synchronization events. We then present three techniques to minimize hypervisor induced overheads on parallel workloads in large-vCPU VMs. The first technique pins application threads to dedicated vCPUs to eliminate inter-vCPU thread migrations, reducing the overhead of emulating inter-processor interrupts (IPIs). The second technique para-virtualizes inter-vCPU TLB flush operations. The third technique enables faster reactivation of idle vCPUs by prioritizing the delivery of rescheduling IPIs. Unlike existing solutions which rely on heavyweight and slow vCPU hotplug mechanisms, our techniques are lightweight and provide more flexibility in migrating large-vCPU VMs. Using several parallel benchmarks, we demonstrate the effectiveness of our prototype implementation in the Linux KVM/QEMU virtualization platform. Specifically, we demonstrate that with our techniques, parallel applications can maintain their performance even when 255 vCPUs are assigned to a VM running on only 6 physical cores.

Bio: Ozgur Ozan Kilic is a Ph.D. candidate at Binghamton University advised by Prof. Kartik Gopalan. His research focuses on virtualization, cloud and high-performance computing.

This event is funded by GSOCS, a subsidiary of GSO, using Student Activity Fee funds

Refreshments will be provided!