Towards Real-Time Clouds for CPS

Chenyang Lu
Cyber-Physical Systems Laboratory
Department of Computer Science and Engineering
http://www.cse.wustl.edu/~lu/
Challenge: Real-Time Clouds

- Support multiple CPS applications on shared resources
- Preserve real-time performance guarantees

- Multi-level real-time resource management
  - Real-time virtualization within a host
  - Distributed resource management within a cloud
  - Real-time networking
Clouds are not real-time today

- Virtualization technology underlying clouds
  - Xen: virtual machine monitor for Amazon EC2
- CPU: proportional-share scheduling

- I/O is worse
  - Vague “performance indicator”: low/medium/large
  - Or pay a lot to get dedicated physical network resources
RT-Xen

- Real-time virtualization platform based on Xen
- Compositional real-time CPU scheduling for VMs
- Real-time communication architecture
- Open-source: http://sites.google.com/site/realtimexen/
RT-Xen: Compositional Scheduling

- Provides temporal isolation and real-time guarantee
- Computes minimum-bandwidth resource models for VMs

Virtual Machine Monitor
- Resource Model
- Scheduler
- Workload

Virtual Machines
- Resource Model
- Scheduler
- Workload

Periodic Resource Model (period, budget)
Rate Monotonic
Periodic Tasks
Component

“Realizing Compositional Scheduling through Virtualization”, Real-Time and Embedded Technology and Application Symposium (RTAS), 2012
Xen Credit vs. Real-Time VM Scheduling

Credit scheduler → poor real-time performance

Real-time VM scheduling helps!

“RT-Xen: Towards Real-Time Hypervisor Scheduling in Xen”,
ACM International Conferences on Embedded Software (EMSOFT), 2011
Towards Real-Time Clouds

- RT-Xen is only a small step
  - Real-time scheduling → virtualization shows promise

- Open challenges
  - Semantics of service guarantees for CPS applications
  - Parallel real-time computing in the cloud
  - Distributed resource management in the cloud
  - Predictable adaptation to dynamic CPS workloads and failures
  - Real-time clouds at a global scale
  - Integration with virtualized real-time networking