Tracing Kernel Virtual Machines (KVM) and Linux Containers (LXC)

Julien Desfossez

École Polytechnique de Montréal
Révolution Linux

25 juin 2010
Plan

1. Introduction
   - Virtualization and Contextualization Technologies

2. Tracing the Hypervisor
   - Full Virtualization Technology

3. Tracing LXC
   - Contextualization Technology

4. Conclusion
KVM

- Mainline since 2.6.20
- Supports Native (Full) Virtualization using Intel VT-x or AMD-V
- Supports Paravirtualization for certain drivers with the VirtIO framework
LXC Linux Containers

- Mainline since 2.6.29
- Userspace containers
- Provide full resource isolation and control
- Replaces OpenVZ, Vservers, FreeVPS
Plan

1. Introduction
   - Virtualization and Contextualization Technologies

2. Tracing the Hypervisor
   - Full Virtualization Technology

3. Tracing LXC
   - Contextualization Technology

4. Conclusion
Role of the Hypervisor

- Virtual Machine Monitor at ring -1
- The host switches between Host State and Guest State using VMCS
- \textit{vmentry} when the CPU is given to the virtual machine (the guest)
- \textit{vmexit} when the CPU is given back to the host (on HLT or exceptions/traps)
Challenges in tracing KVM

- The host is "blind" between the \textit{vmentry} and \textit{vmexit}
- The hypervisor controls the TSC:
  - On exit: save TSC
  - Before re-entrance: take delta + exit overhead
  - Subtract from TSC offset

$\rightarrow$ Traces recorded in the host and in the guest are not directly aligned
Objectives in tracing KVM

- Monitoring the whole system (host and guest(s))
- Presenting graphically the interaction between the host and the guest(s)
- Integrating with the dependency analysis
  ➤ Need to add new CPU states
Plan

1. Introduction
   - Virtualization and Contextualization Technologies

2. Tracing the Hypervisor
   - Full Virtualization Technology

3. Tracing LXC
   - Contextualization Technology

4. Conclusion
LXC

- No virtualization involved
- Adds the notion of context to process groups
- Allow resources control
Challenges in tracing LXC

- Need to add the context attribute in the trace
- Graphically represent each container as a separate resource
- Identify the interactions between different groups
- Identify the impact of resources control
# Plan

1. **Introduction**
   - Virtualization and Contextualization Technologies

2. **Tracing the Hypervisor**
   - Full Virtualization Technology

3. **Tracing LXC**
   - Contextualization Technology

4. **Conclusion**
Conclusion

- Tracing two types of isolation technologies:
  - KVM: full virtualization with potentially no guest assistance
  - LXC: contextualization technology
- Synchronizing the traces between the host and guest(s)
- Visually representing the virtualized CPU state