Inception: Towards a Nested Cloud Architecture

Changbin Liu, Yun Mao
AT&T Labs – Research
HotCloud’13
Motivation

• IaaS cloud is good!
• But
  – VM live migration?
  – Intrusion detection [OSDI’02], VM replication [VLDB’11], VM instant spawning [EuroSys’11]...?
  – Build your own large-scale cloud (OpenStack, CloudStack)?
Motivation

• Now that we have VMs from cloud providers, why not use them as “physical” machines?
• Idea: nested cloud
Inception Cloud

• Hypervisor flexibility
  – VM live migration, replication, instant spawning, intrusion detection ...

• App-specific resource allocation
  – Customize subscription ratio, VM consolidation

• Low operation overhead
  – No physical metals, no cooling, power, cabling, rack design, malfunctioned hardware replacement, data center real estate

• Multiple providers
  – Private + public clouds over geographically wide area

© 2013 AT&T Intellectual Property. All rights reserved. AT&T and the AT&T logo are trademarks of AT&T Intellectual Property.
Application Scenarios

- Web hosting: instant VM spawning, replication, intrusion detection
- Continuous integration: comprehensive test, full network, cleanup
- Virtual desktop: VM live migration, consolidation, virus scanner

None is achievable in current IaaS clouds
Design and Implementation

• Cloud management software: stock OpenStack
• Hypervisor: KVM nested virtualization

**Network**

– Reality VM (rVM): one private IP + one public IP (optional)
– Inception VM (iVM): layer-2 networking, easy to configure and manage, iVM live migration
Network Design Overview

eth1: private

eth0: management

eth2: public

- Virtualized layer-2 networking
  - Open vSwitch + VXLAN (UDP-based L2 over L3 tunneling)
  - Isolated with each other, transparent to iVMs

© 2013 AT&T Intellectual Property. All rights reserved. AT&T and the AT&T logo are trademarks of AT&T Intellectual Property.
Private Network

• iVMs do not attach to obr1
  – obr1 is out of band for OpenStack network controller
  – Incompatible with Linux bridge based networks
  – Open vSwitch does not well support iptables/ebtables
Public Network

- Floating IP: NAT between public and private IPs

Each rVM can host many iVMs with any number of floating IPs
- Public/private traffic is not disrupted during iVM live migration
- Support “recursively” nested inception clouds
Prototype

• OpenStack core components: compute, storage, network, etc
• Chef automation: both OpenStack and Open vSwitch
• Source code: [github.com/stackforge/inception](https://github.com/stackforge/inception)
• Deployment on our private cloud
• Launch an inception cloud in ~20 minutes
On-going Work

• Performance
  – CPU and memory: 5-10% performance overhead
  – Network and storage: significant performance degradation

• Scalable virtualized layer-2 networking
  – Target: span across multiple data centers, each with ~10,000 iVMs
  – SDN-based control plane: ARP, DHCP, live migration, security group, broadcast, multicast, etc

• OpenStack
  – Facilitate continuous integration
  – New hypervisor features, resource scheduling/optimization
“An idea is like a virus. Resilient. Highly contagious. And even the smallest seed of an idea can grow. It can grow to define or destroy you.”
Thank you!

Q&A

github.com/stackforge/inception