## BINGHAMTON UNIVERSITY

State University of New York

Span VM: Multi-Hypervisor Virtual Machines Enabling An Ecosystem of Hypervisor-Level Services In Cloud PI: Kartik Gopalan, Binghamton University (SUNY), kartik@binghamton.edu



Problem: Support for 3<sup>rd</sup>-Party Hypervisor-level Services

## Solution: Compartmentalize Services & Share Guest Control

**1. Growing Number of Hypervisor**level Services: VM Introspection, Intrusion Detection, High Availability, Live Migration, Live Patching, etc.



2. Guests Cannot Simultaneously Use Multiple 3<sup>rd</sup>-party Services: E.g. Cross-cloud migration, Customized guest security, Attestation, etc.

**Featurevisors** (F) : 3rd-party deprivileged "Hypervisors" providing guest services. **Hyperplexor** : Base L0 hypervisor.

Approach: Transparent and Simultaneous Control of Guest by Multiple L1 Hypervisors







- **Guest Transparent:** No modifications to guest.
- Attach/Detach L1s to/from guest at runtime: Partial/full control over guest memory, VCPUs, and I/O devices.
- **Event Subscription:** L1s subscribe to guest events via L0.

nested@spanvm-lla\$ sudo tcpdump -q -i br0 -n src 10.128.24.1 tcpdump: verbose output suppressed, use -v or -vv for full protocol decode istening on br0, link-type EN10MB (Ethernet), capture size 96 bytes 17:29:31.716554 ARP, Request who-has 10.128.0.1 tell 10.128.24.1, length 28 17:29:43.824093 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 0 7:29:43.829140 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 0 17:29:43.846370 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 32 17:29:43.848073 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 0

## Status, Results, and Future Work

**Key Publications:** 

Guest infected

l2g@l2g:~\$ ./evil &

l2g@l2g:~\$ ps -e | grep evil l2g@l2g:~\$

[1] 883

- Multi-hypervisor Virtual Machines: Enabling an Eco-system of Hypervisor-level Services, Accepted in USENIX ATC, 2017
- Enabling Hypervisor-as-a-service Clouds with Ephemeral Virtualization, VEE 2016.

nected@cnenym 11h	¢	1 1	D = O
₽	nested@spanvm-l1b: ~/volatility-2.4 66x13		l2g@l2g:~29x12
17:29:44.015604	IP 10.128.24.1.22 > 10.128.	0.9.48050: tcp 56	
1/:29:44.014/00	1P  10.128.24.1.22 > 10.128.	0.9.48050: tcp 0	
17.20.44 014700			
17:29:44.013728	IP 10.128.24.1.22 > 10.128.	0.9.48050: tcp 0	
17.23.45.007750	11 10.120.24.1.22 > 10.120.	0.5.40050. CCp 200	
17.29.43 867730	TP 10 128 24 1 22 $>$ 10 128	$0 \ 9 \ 48050 \cdot t cn \ 280$	Era. Hotwork morntoring
1/:29:43.8494/5	1P 10.128.24.1.22 > 10.128.	0.9.48050: tcp 952	1 1a: Network Monitoring
		() $()$ $()$ $()$ $()$ $()$ $()$ $()$	

l2g@l2g:~\$ cat evil.c main(void) ested@spanvm-l1b\$ python vol.py -f /mnt/l2dump --profile=LinuxUbu tu1204x64 plugin name linux psaux | tac | grep evil Volatility Foundation Volatility Framework 2.4 while(1) sleep(1000); with KBeast 1000 1000 ./evi

lested@spanvm-l1b\$ lested@spanvm-l1b\$ nested@spanvm-l1b\$ lested@spanvm-l1b\$ lested@spanvm-l1b\$ nested@spanvm-l1b\$ lested@spanvm-l1b\$

Funded by NSF 1527338: CSR: Small: Multi-hypervisor Virtual Machines -Enabling an Ecosystem of Hypervisors in the Cloud

L1b: Volatility

- **Prototype on KVM/QEMU Platform** 
  - 0—15% overhead on benchmarks: Kernbench, iperf, quicksort.
  - Ephemeral virtualization: 80ms average switching times
  - Page fault servicing: 3.6—4.2us; Event Redirection: 13-41us.
- **Ongoing/Future Work:** 
  - Supporting unmodified L1 hypervisors.
  - Live hypervisor patching.
  - Support on public clouds.

Students: Hardik Bagdi, Rohith Raghavendra, Yaohui Hu, Spoorti Doddamani Collaborators: Dan Williams, Nilton Bila, Umesh Deshpande, IBM Research Labs