

Muti-Hypervisor Virtual Machines

Yaohui Hu, Rohit K. Raghavendra, Hardik Bagdi, Piush Sinha, ^bUmesh Deshpande,

Kartik Gopalan, ^aDan Williams, ^aNilton Bila

Binghamton University, ^aIBM T.J. Watson Research Center, ^bIBM Almaden Research Center

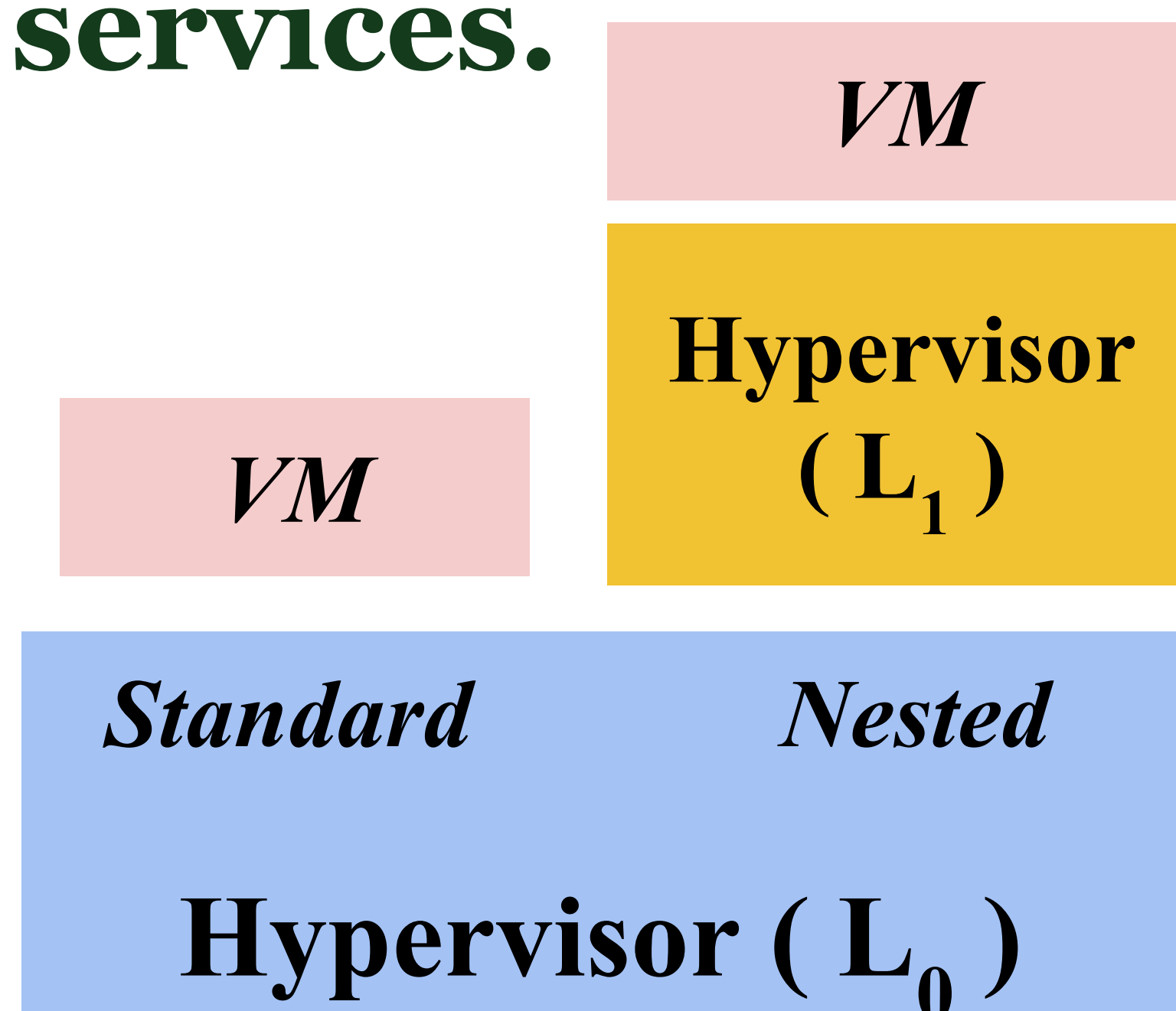
Web: <http://spanvm.github.io/>

Email: spanvm@binghamton.edu

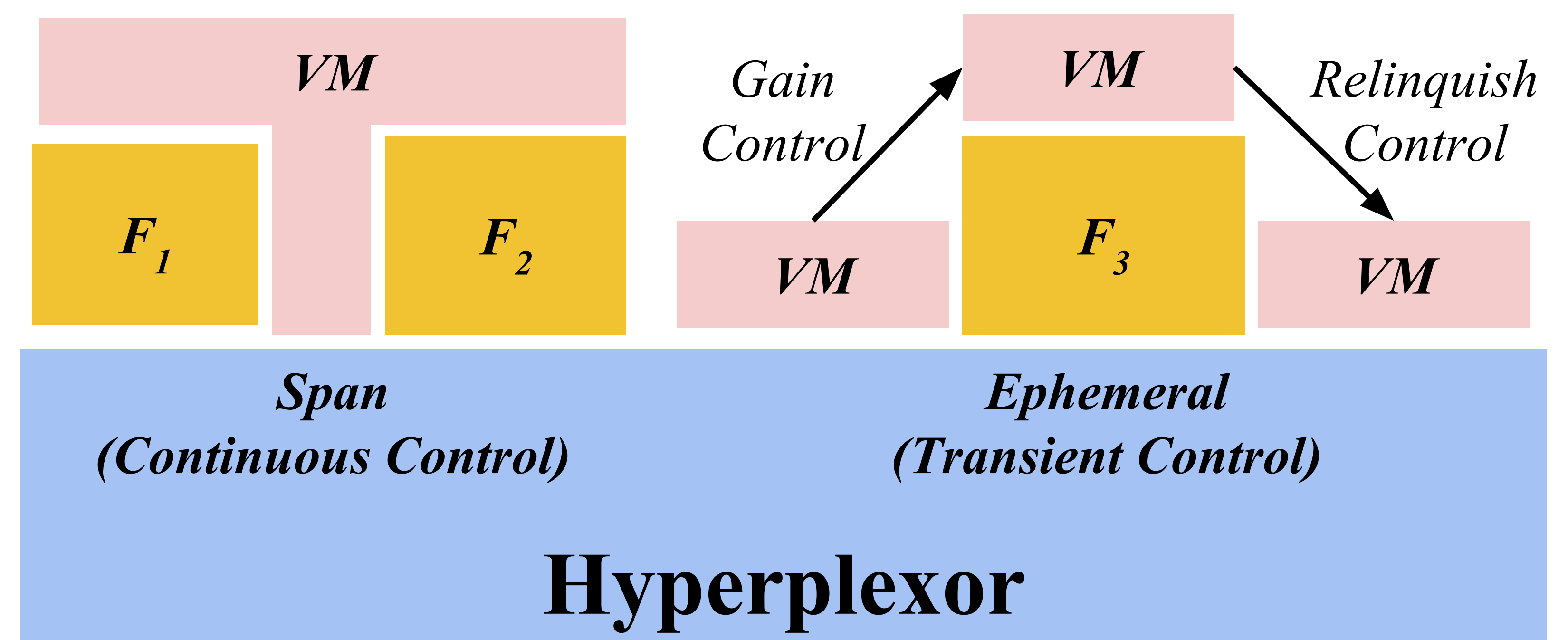
Hypervisors are Getting Too Big

Problem: Too many services.

- VM Introspection
- Intrusion Detection
- High Availability
- Live Migration
- Live Patching
- File systems etc.



Solution: Break Them Up



Featurevisors (F) : L1 “Hypervisors” that provide add-on services.

Hyperplexor : L0 hypervisor that multiplexes hardware for L1 guests.

Challenge: Transparent Control of Guest By Multiple Hypervisors

Guest Memory

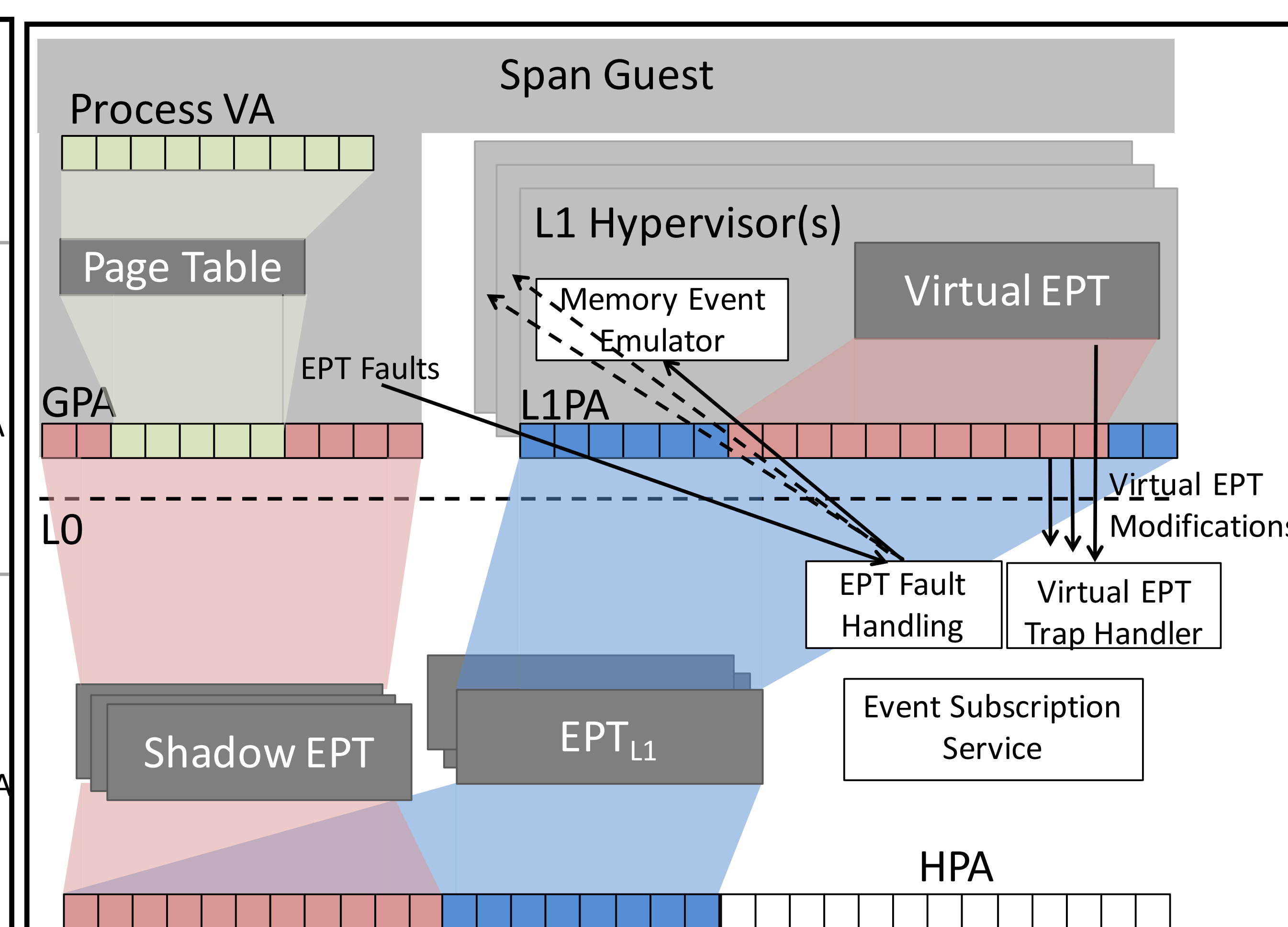
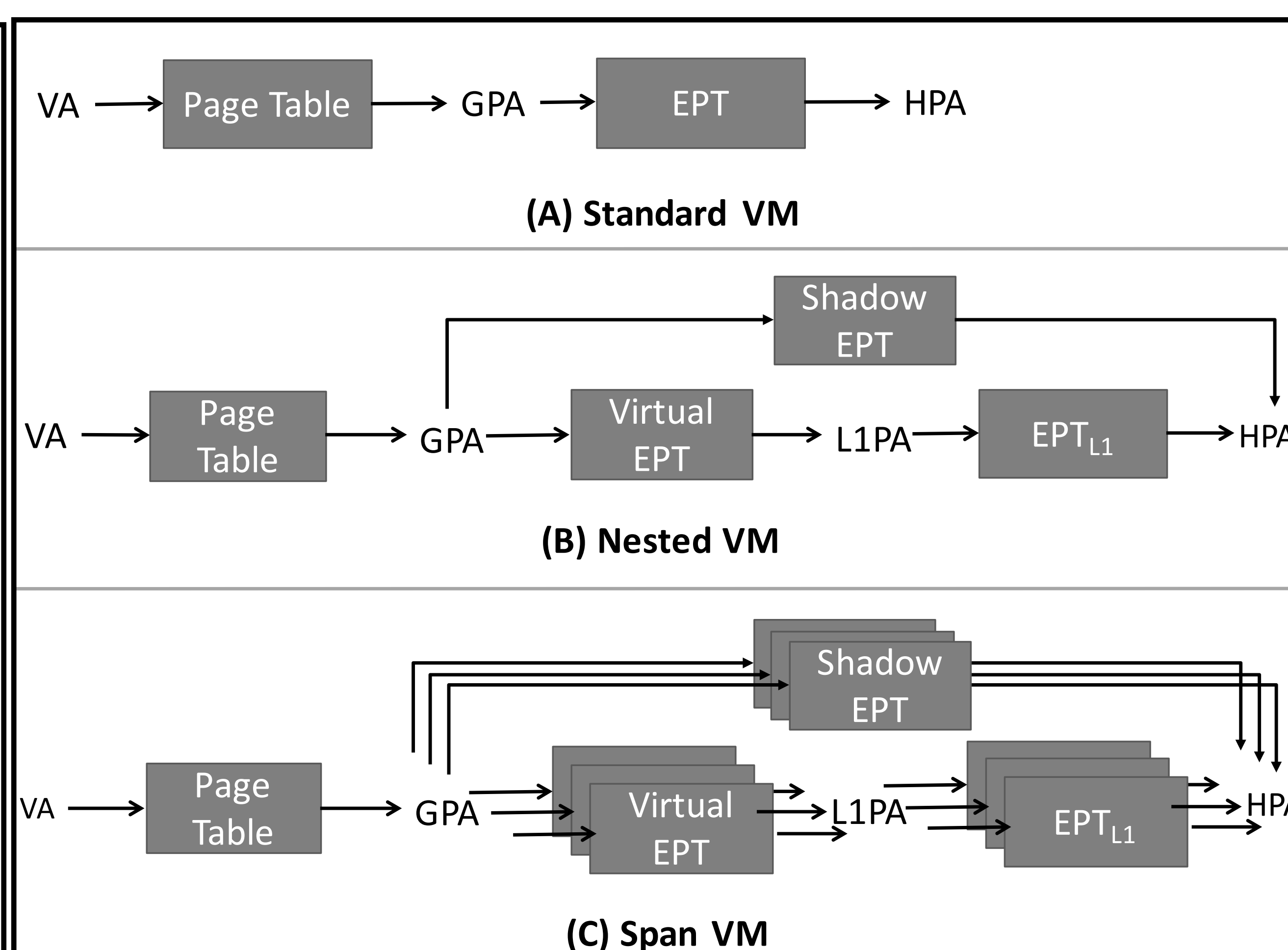
- Shared

Virtual I/O Devices

- Distributed.

VCPUs

- One hypervisor.



Demonstration

```

nested@spanvm-l1a$ sudo tcpdump -q -i br0 -n src 10.128.24.1
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening 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
17:29:43.829140 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
17:29:43.849475 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 952
17:29:43.867730 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 280
17:29:44.013728 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 0
17:29:44.014700 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 0
17:29:44.015604 IP 10.128.24.1.22 > 10.128.0.9.48050: tcp 56

nested@spanvm-l1b$ python vol.py -f /mnt/l2dump --profile=LinuxUbuntu
ntu1204x64 plugin name linux_psaux | tac | grep evil
Volatility Foundation Volatility Framework 2.4
883 1000 1000 ./evil

nested@spanvm-l1b$
nested@spanvm-l1b$
nested@spanvm-l1b$
nested@spanvm-l1b$
nested@spanvm-l1b$
nested@spanvm-l1b$
```

L1a: Network Monitoring

L1b: Volatility

Guest infected with KBeast

Performance

- On modified KVM/QEMU platform
- Common benchmarks
 - Kernbench, iperf, Quicksort
- 0–15% overhead vs. standard VMs
- Ephemeral Virtualization
 - 80ms average switching times.
- Page-fault servicing: 3.6–4.2μs
- Interrupt redirection: 13–41μs