

Multi-Hypervisor Nested Virtual Machines

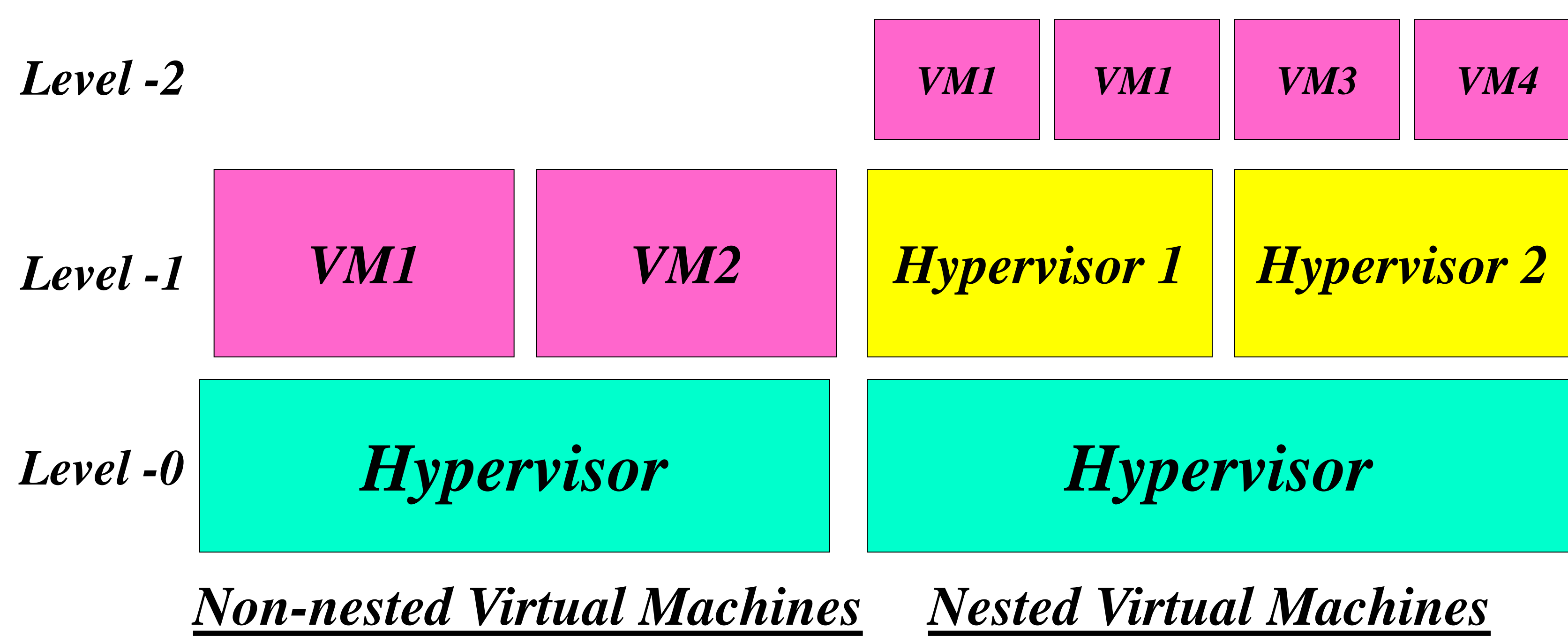
Yaohui Hu, Siddhesh Phadke, Kartik Gopalan, Michael R.Hines*

Computer Science, Binghamton University

*IBM Research Labs, Yorktown Heights

Contacts: yhu15@binghamton.edu, kartik@binghamton.edu

Traditional Nested Virtualization



Benefits

- ❖ Users can run their own hypervisors in IaaS clouds.
- ❖ Live migration of hypervisors + guest VMs as a single entity.
- ❖ Hypervisor-level intrusion detection/prevention.

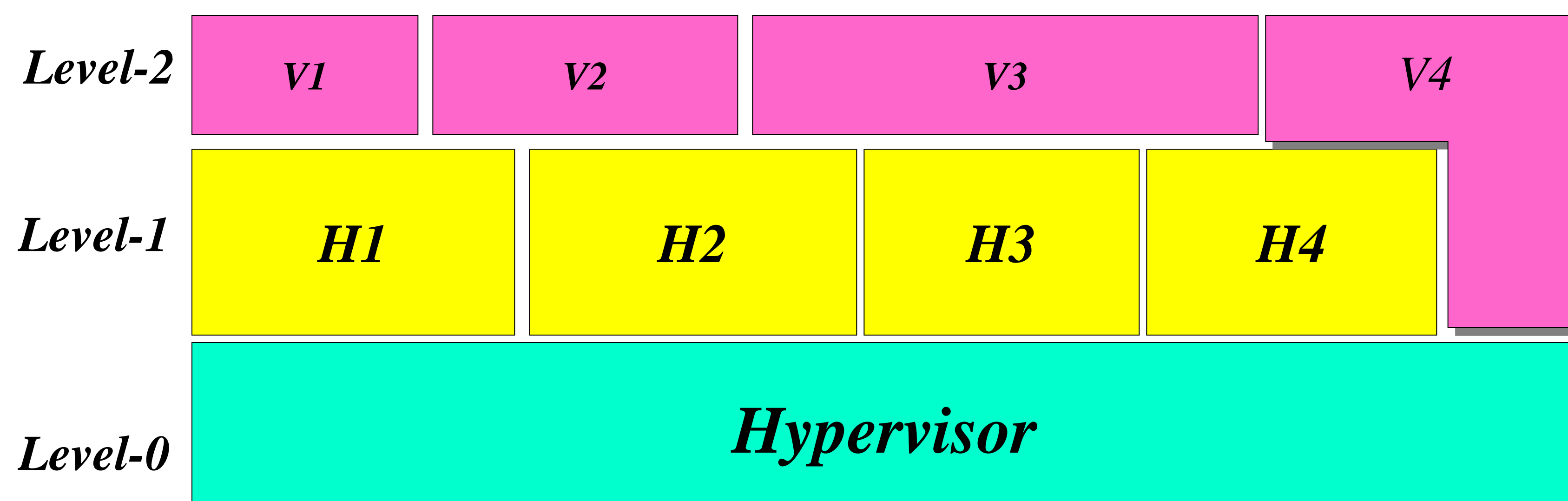
Limitation

- Presently, an L2 VM can run on only one L1 hypervisor at a time.
 - L2 VM cannot run on multiple co-located L1 hypervisors.
- E.g: An L2 VM cannot simultaneously run on a commodity L1 hypervisor and another L1 hypervisor providing intrusion-detection.

Multi-Hypervisor Nested Virtualization

Objective

- Develop systems support to run unmodified Nested VMs simultaneously on multiple hypervisors.

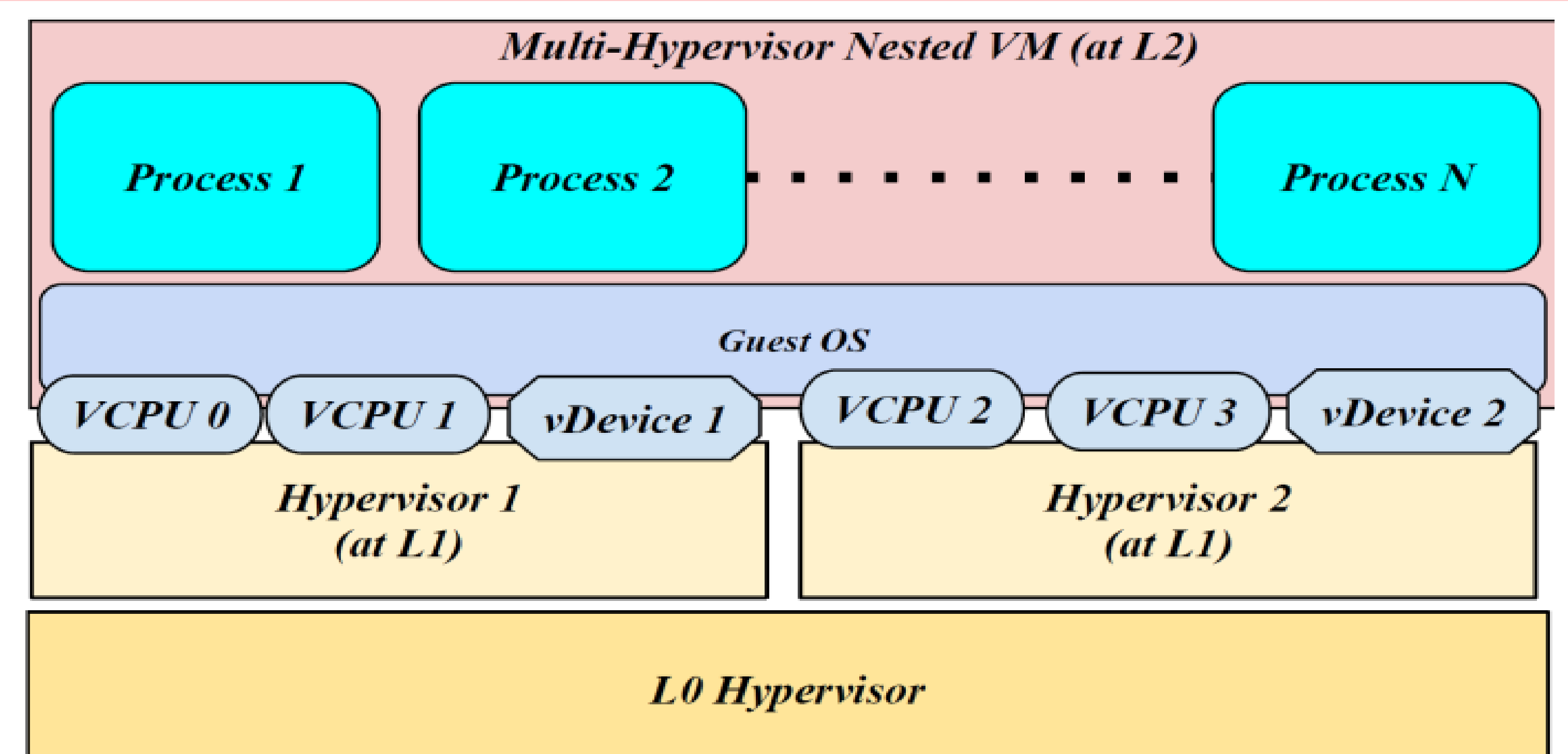


- ❖ $H1, H2, H3,$ and $H4$ are Level-1 hypervisors.
- ❖ $V1, V2, V3,$ and $V4$ are Level-2 VMs.
- ❖ $V1$ runs only on $H1$ as a traditional nested VM.
- ❖ $V2$ runs on $H1$ and $H2$.
- ❖ $V3$ runs on $H2, H3$ and $H4$.
- ❖ $V4$ runs on $H4,$ and $L0$.

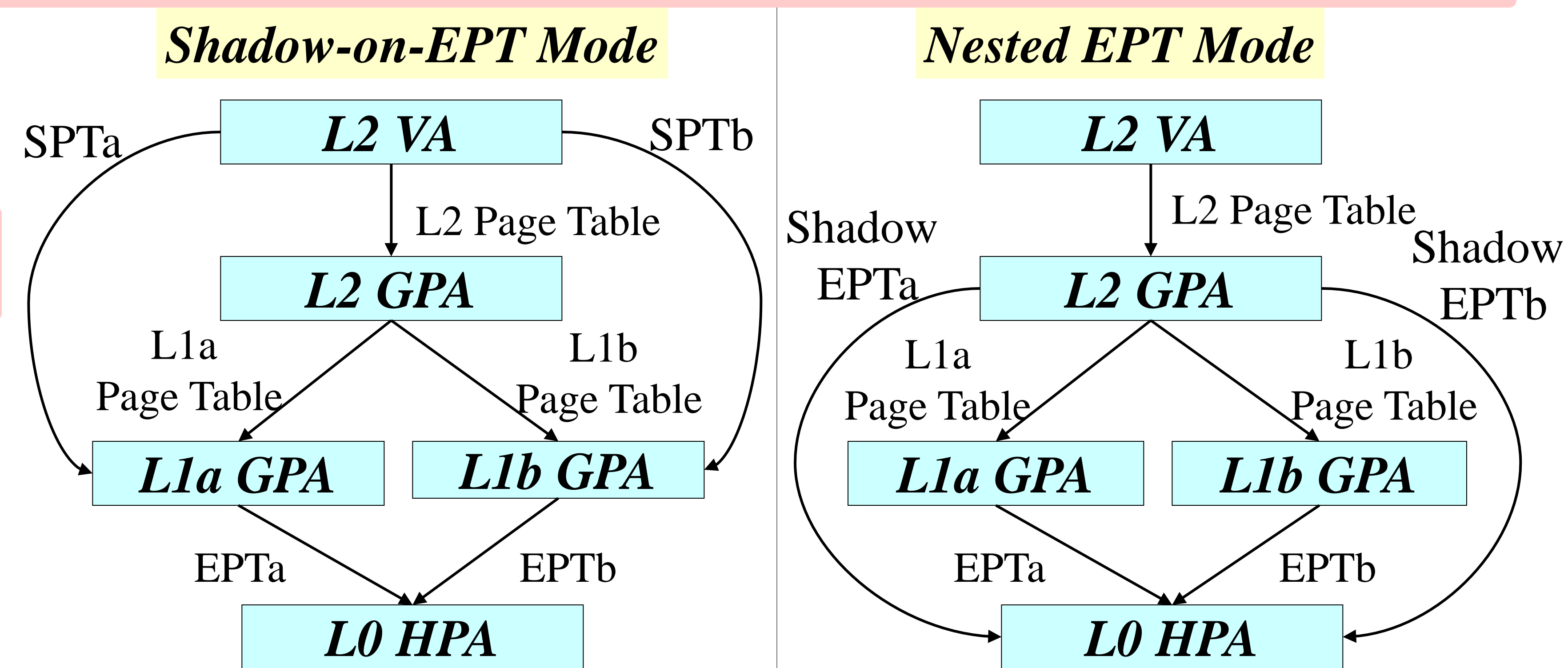
Challenges

1. Sharing the L2 memory across multiple L1s
 - o Two modes: Shadow-on-EPT and Nested EPT.
2. Distributing L2 vCPUs across multiple L1s
3. Forwarding inter-processor interrupts (IPIs) across L1s
4. Distributing I/O using virtio or direct device assignment.

Resource Distribution



Memory Mapping



Prototype and Performance

Current Status

- Developed prototype to run L2 VM on two L1 hypervisors running KVM/QEMU (as in V2)
- Memory mapping uses shadow-on-EPT mode.
 - Ongoing work on Nested EPT mode
- Virtio-based I/O distribution.
 - Ongoing work on direct assignment.

	Kernbench			
	Host	Guest	Nested	Span
Run time (sec)	136.15	146.31	634.70	674.79
STD dev.	8.09	1.13	8.79	9.68
% overhead vs. host	-	7.5	366.2	395.6
% overhead vs. guest	-	-	333.8	361.2
% overhead vs. nested	-	-	-	6.3
%CPU	97	90	100	100

	netperf			
	Host	Guest	Nested	Span
Throughput (Mbps)	940.5	930.17	343.92	311.36
STD dev.	0.38	0.64	26.12	12.82
% degradation vs. host	-	1.1	63.4	66.9
% degradation vs. guest	-	-	63.3	66.5
% degradation vs. nested	-	-	-	9.5