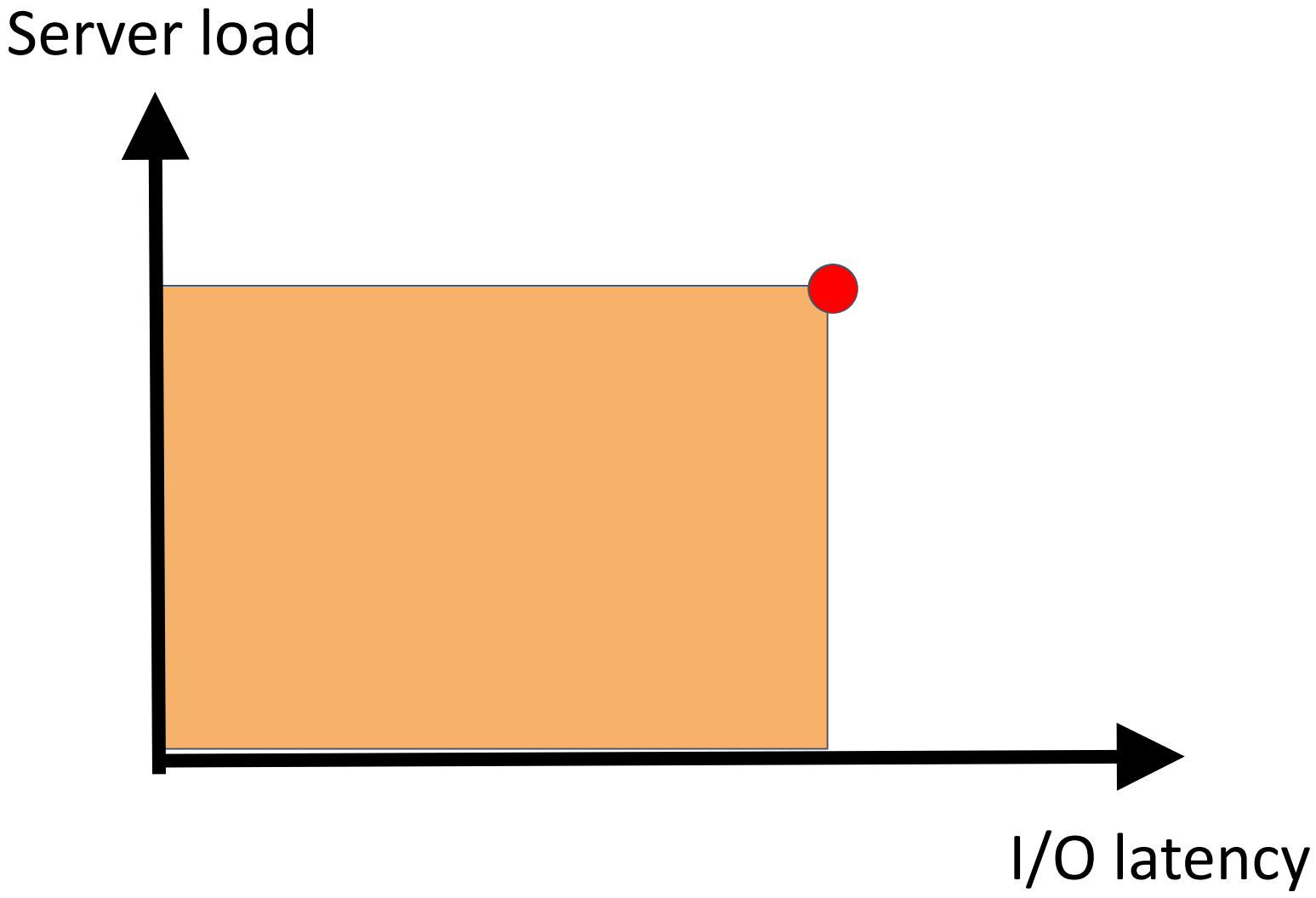
Understanding the Impact of Cache Locations on Storage Performance and Energy Consumption of Virtualization Systems

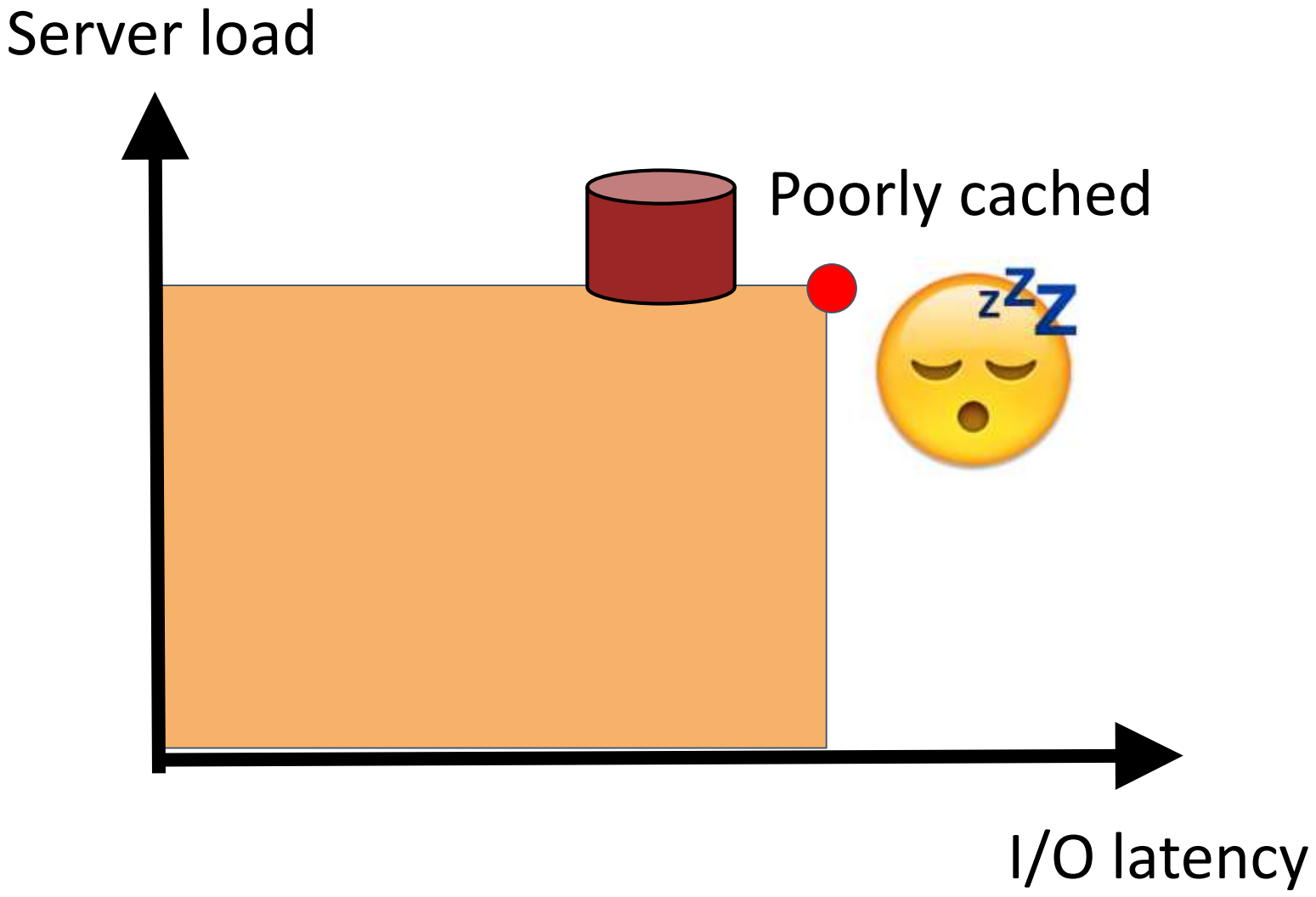
Tao Lu, Ping Huang, and Xubin He
Virginia Commonwealth University

Ming Zhang
EMC Corporation

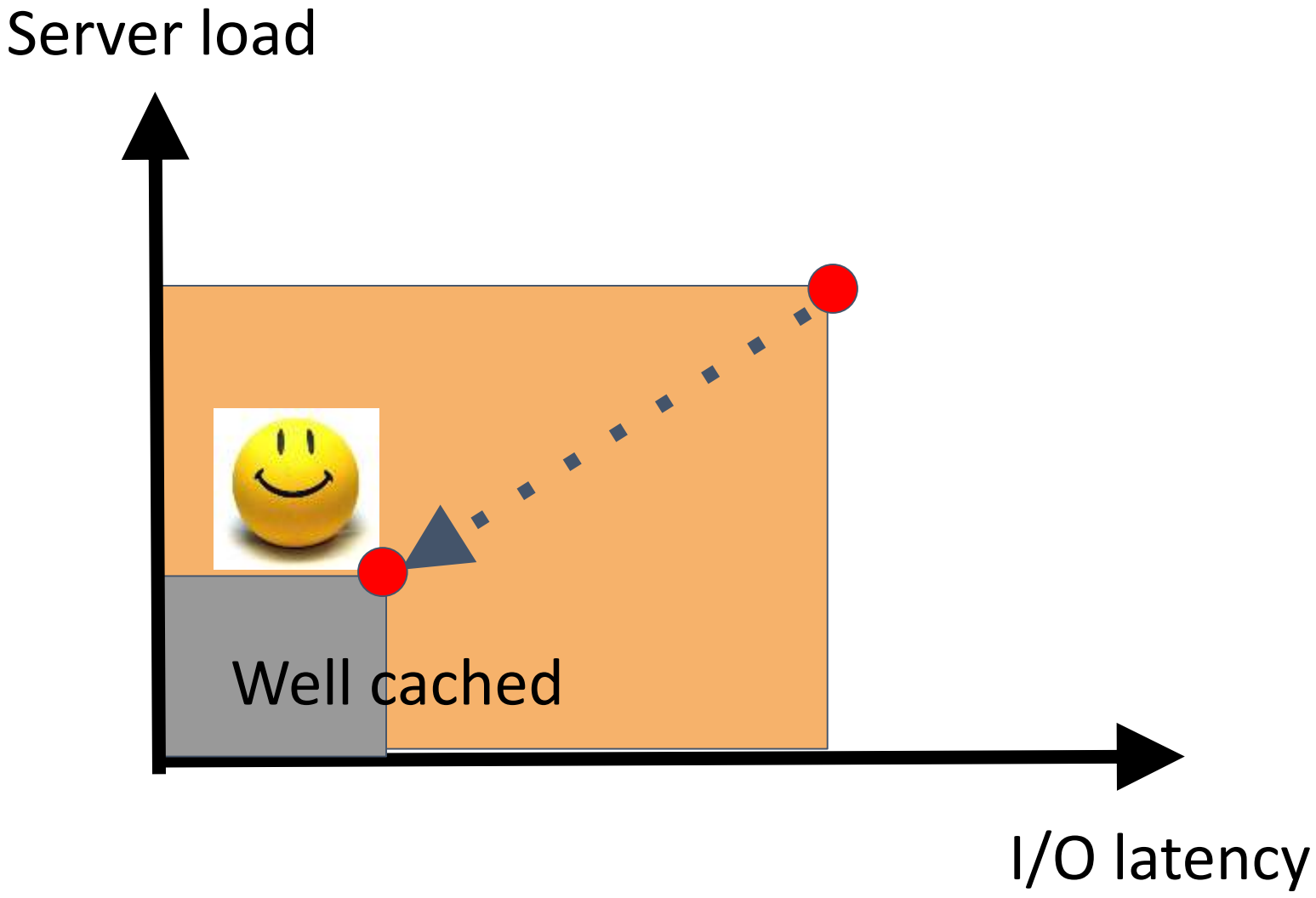
Caching: one key frees two birds



Caching: one key frees two birds

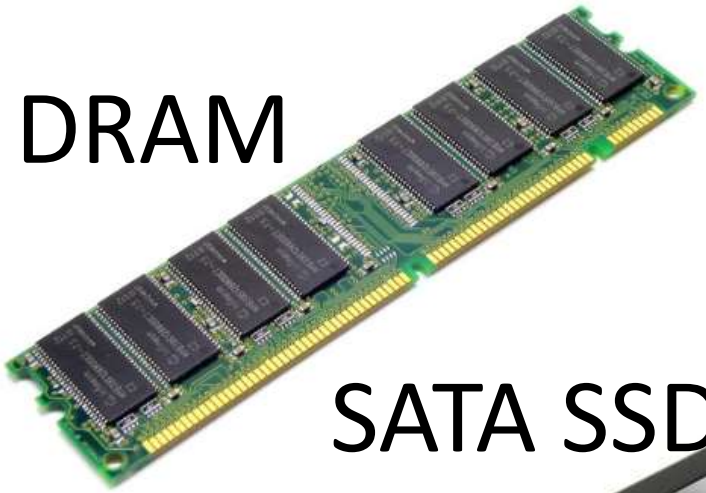


Caching: one key frees two birds



Cache devices

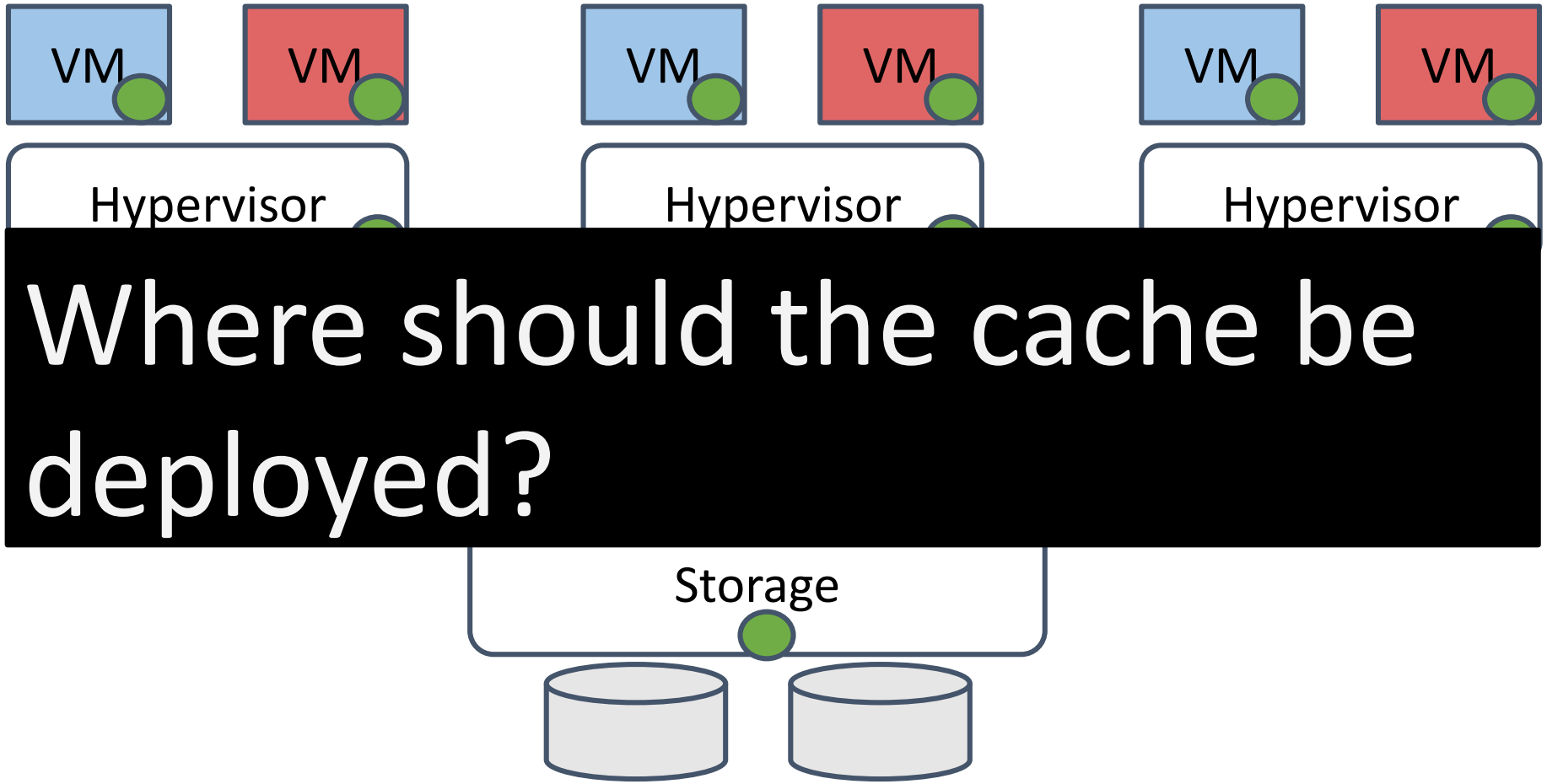
DRAM



PCIe SSD

SATA SSD

Cloud caching: a puzzle



Cloud caching: traditional wisdom

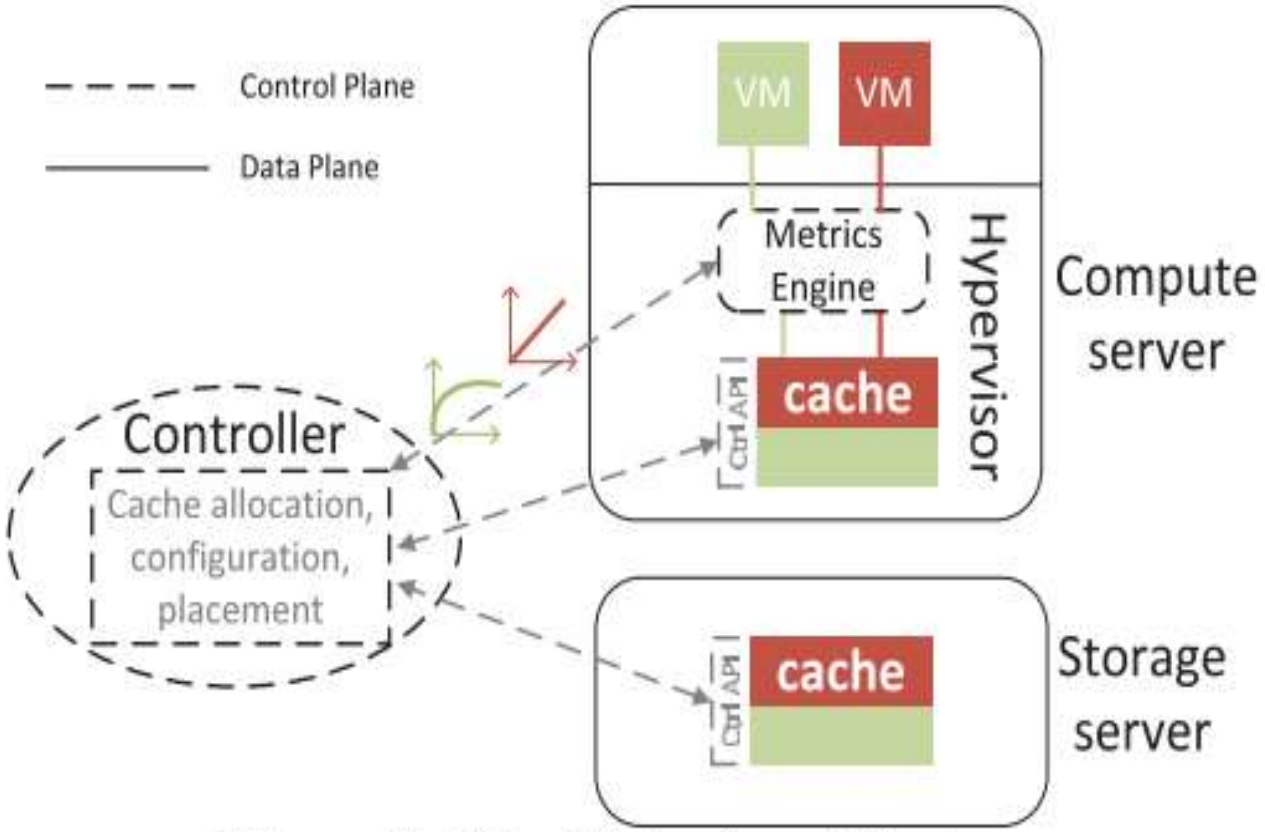


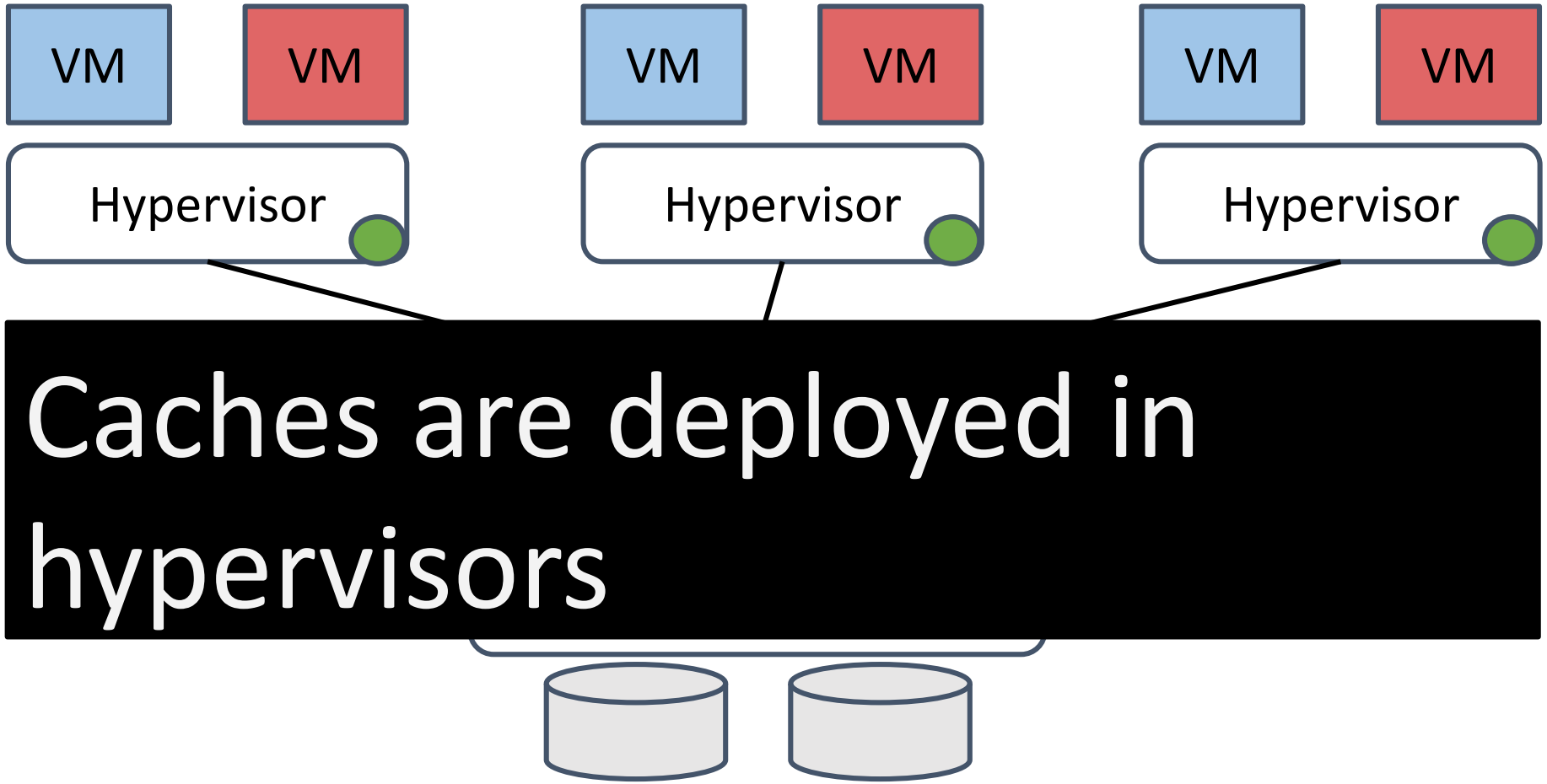
Figure 3: The Moirai architecture.

Ioan Stefanovici et. al., Software-Defined Caching: Managing Caches in Multi-Tenant Data Centers, SoCC'15

Cloud caching: traditional wisdom (cont'd)

- D. Arteaga et. al., ***Cloudcache: On-demand Flash Cache Management for Cloud Computing*** , FAST'16
- Hwang, J. et.al., ***Mortar: Filling the Gaps in Data Center Memory***, *VEE'14*
- Byan, S. et.al., ***Mercury : Host-side Flash Caching for the Data Center***, MSST'12

Cloud caching: traditional wisdom



Is hypervisor the best location to deploy caches?

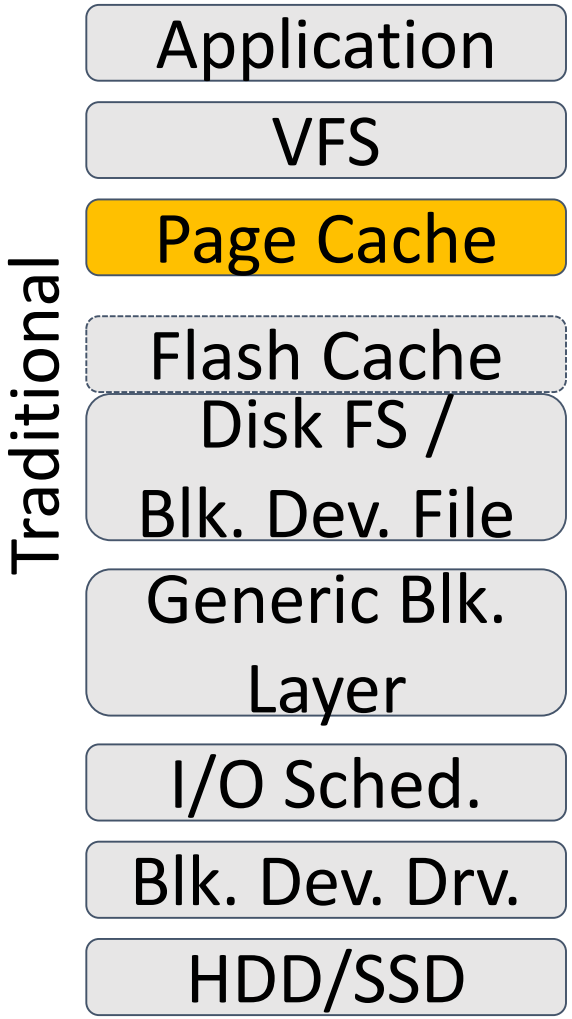
Overview

- Cloud storage stacks
- Impact of cache locations on system performance and energy efficiency
- Potential optimizations

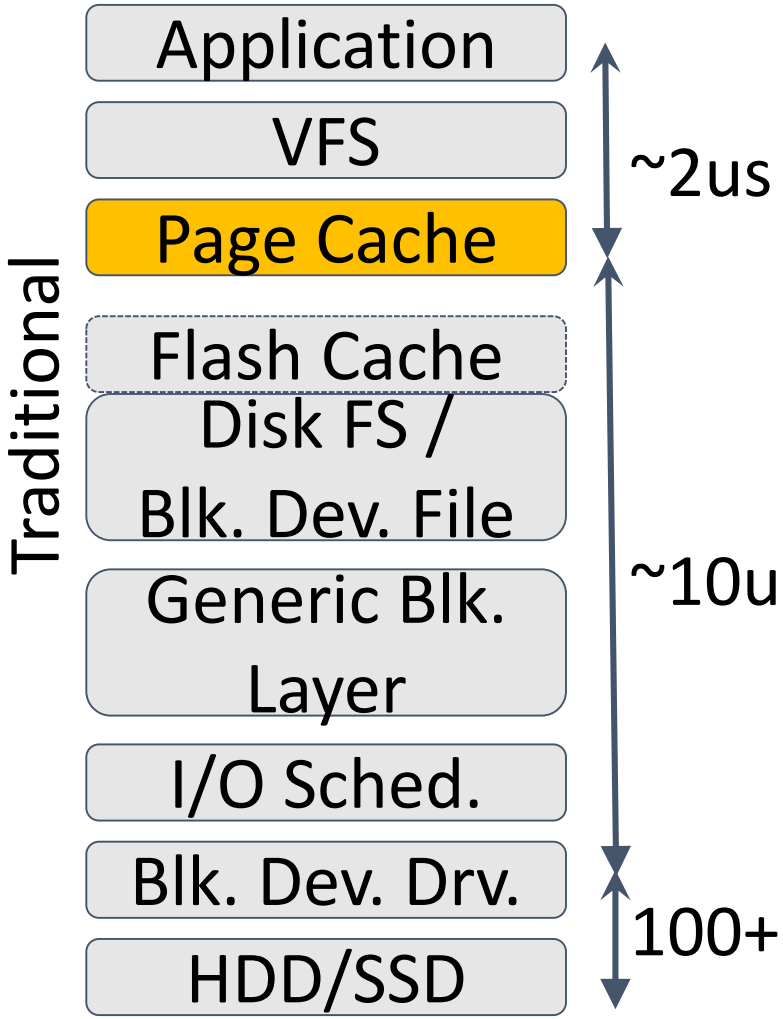
Overview

- Cloud storage stacks
- Impact of cache locations on system performance and energy efficiency
- Potential optimizations

Storage stacks



Storage stacks



Run *fio* micro-benchmark on a physical machine

- Run *fio* once to warm up the page cache
- Run *fio* for a second time to measure the page cache performance

Run *fio* on a physical machine

- Build a RAMDisk
- Run *fio* on the RAMDisk with O_DIRECT enabled to bypass the page cache

Run *fio* on a physical machine

- Run *fio* on a SSD/HDD with O_DIRECT enabled to bypass the page cache

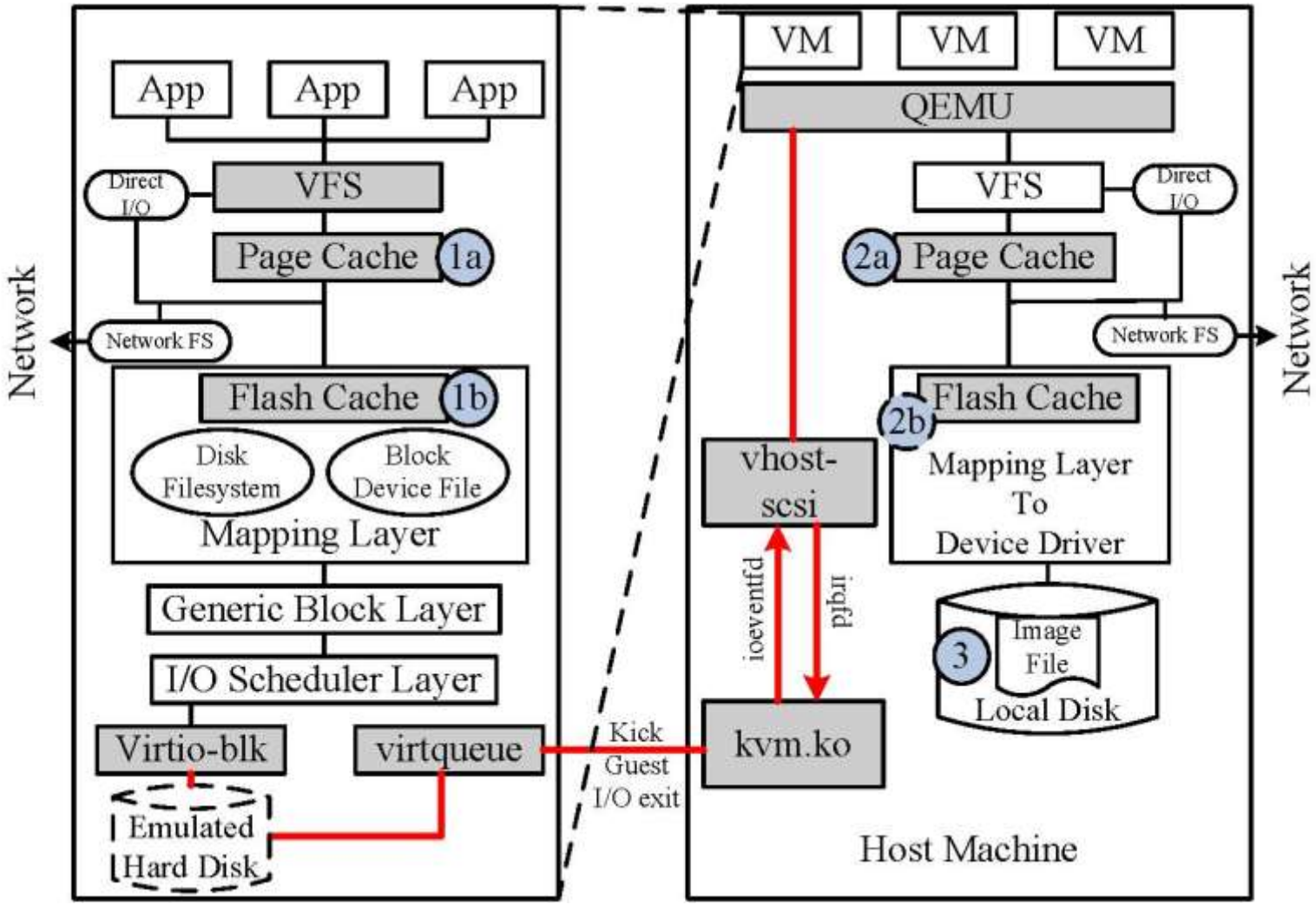
Storage stacks

Traditional

Application
VFS
Page Cache
Flash Cache
Disk FS / Blk. Dev. File
Generic Blk. Layer
I/O Sched.
Blk. Dev. Drv.
HDD/SSD

~2us
~10us
100+ us

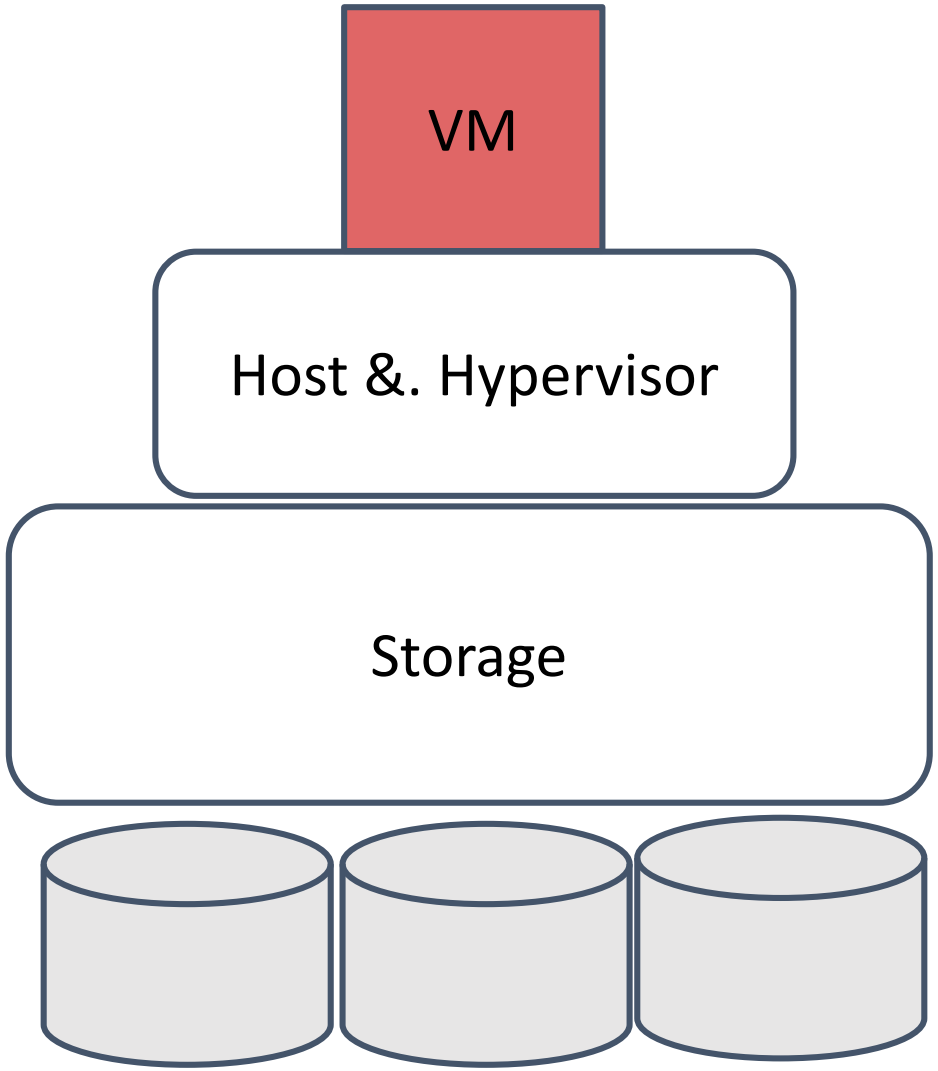
Virtualization



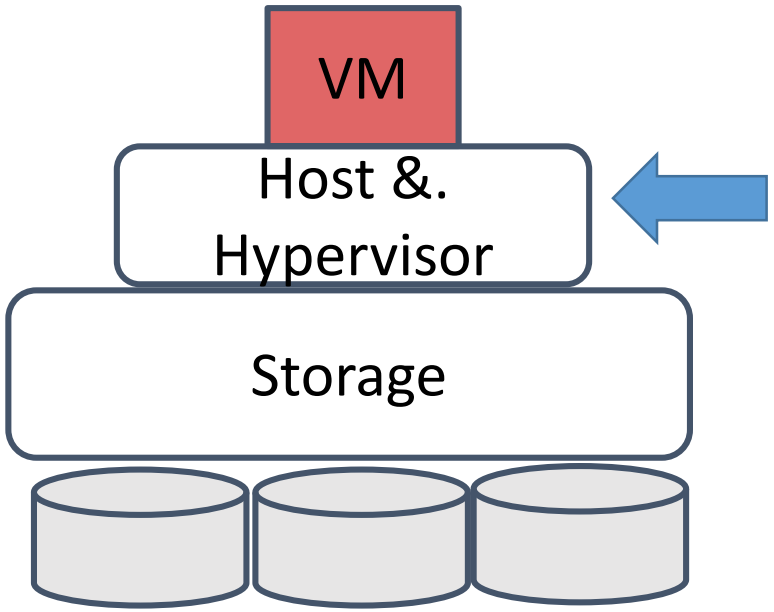
Overview

- Cloud storage stacks
- Impact of cache locations on system performance and energy efficiency
- Potential optimizations

Testbed

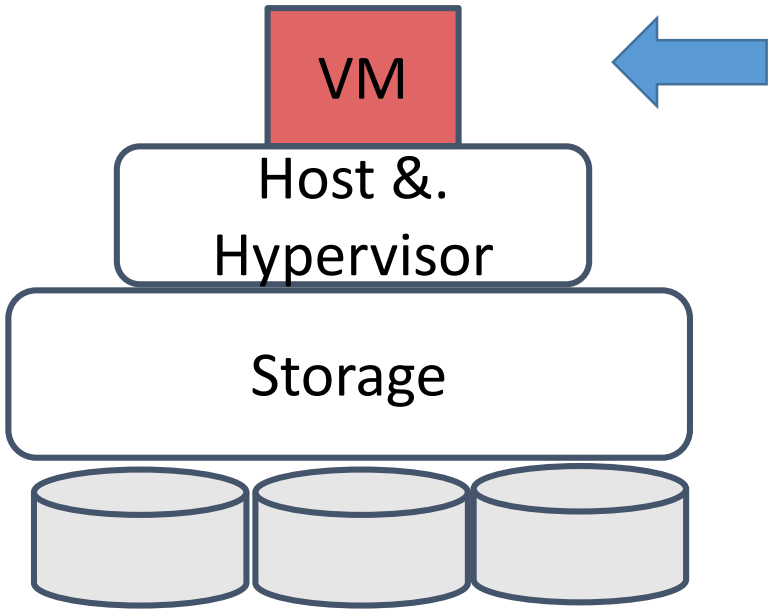


Testbed



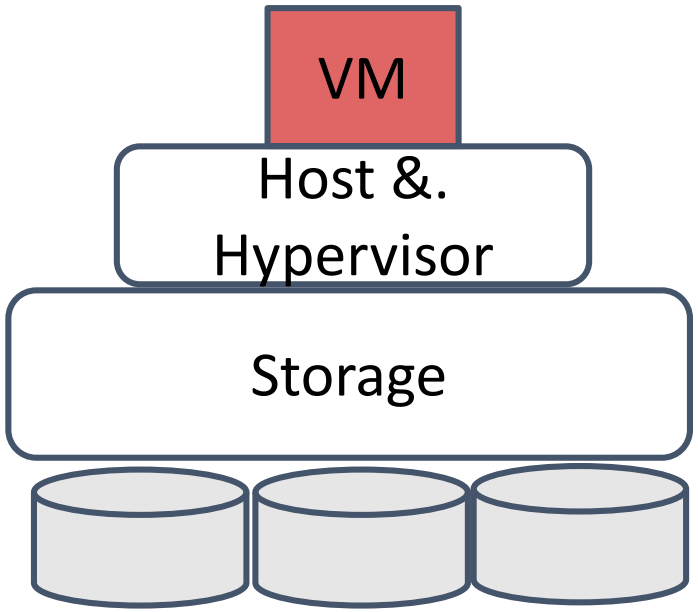
- Physical machine and Host OS
  - AMD Phenom II X4 B95 Quad-core 3.0 GHz processor with AMD-V virtualization support
  - 8GB DRAM
  - Host OS is a 64-bit Ubuntu 15.04 with Linux kernel version 3.19.0- 30-generic.
  - QEMU emulator version 2.4.1 and KVM are used as the hypervisor.

Testbed



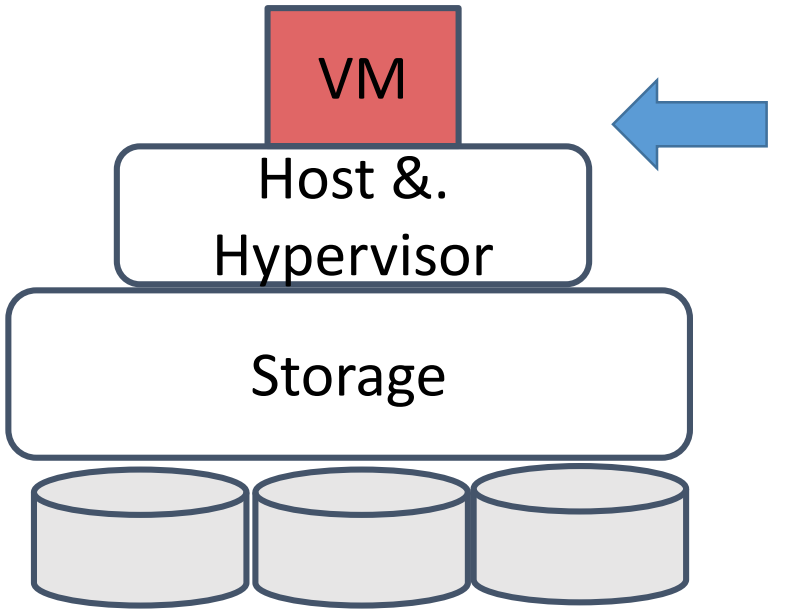
- VM & Guest OS
  - 2 VCPUs and 2GB memory
  - Ubuntu 15.04 64-bit Server Cloud Image as the guest OS

Testbed



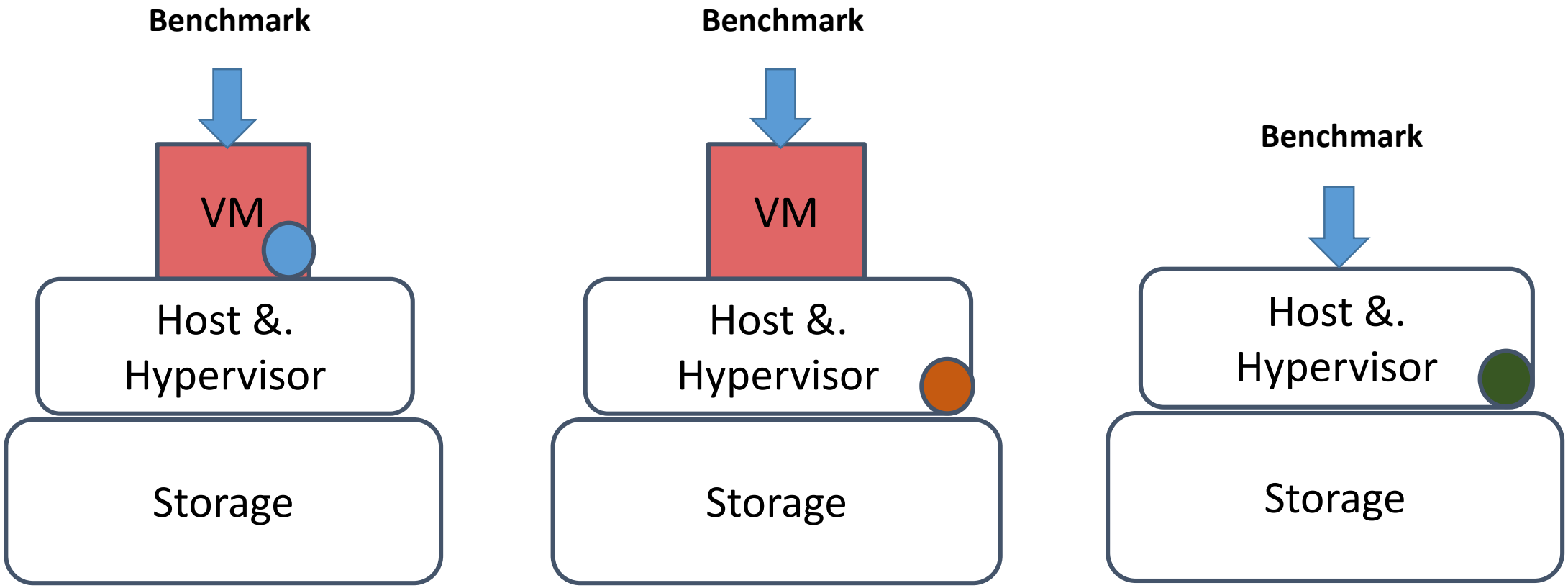
- Disks
  - One 500GB Samsung 850 Pro SATA SSD
  - One 240GB OCZ RevoDrive 3 PCIe SSD

Testbed

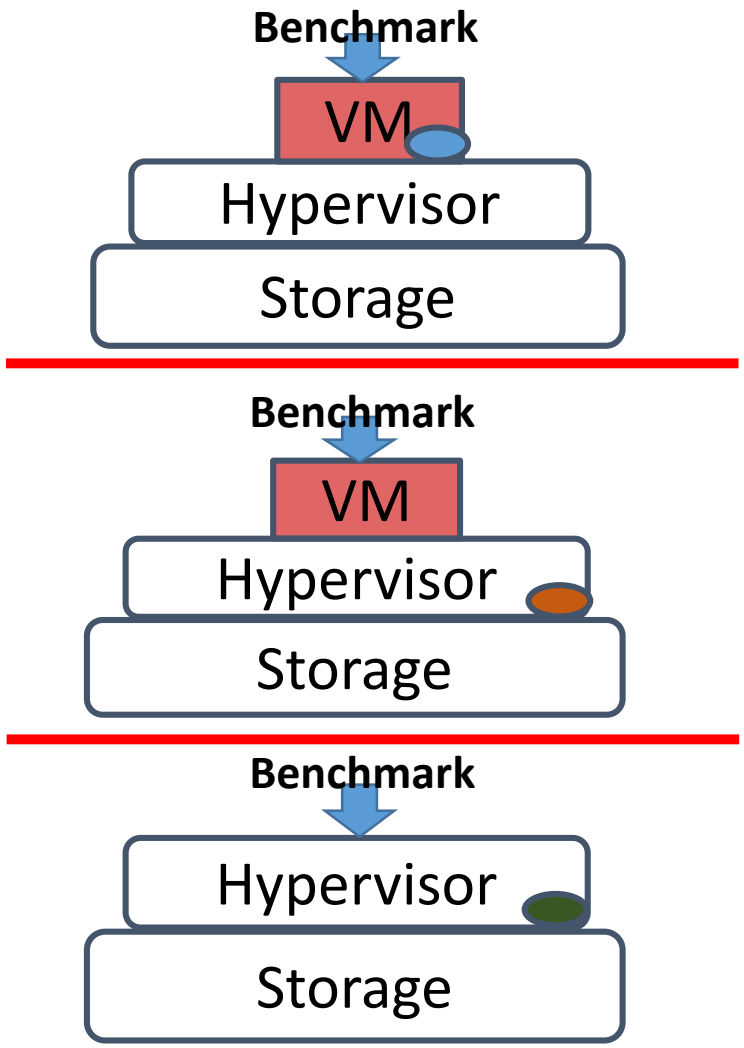


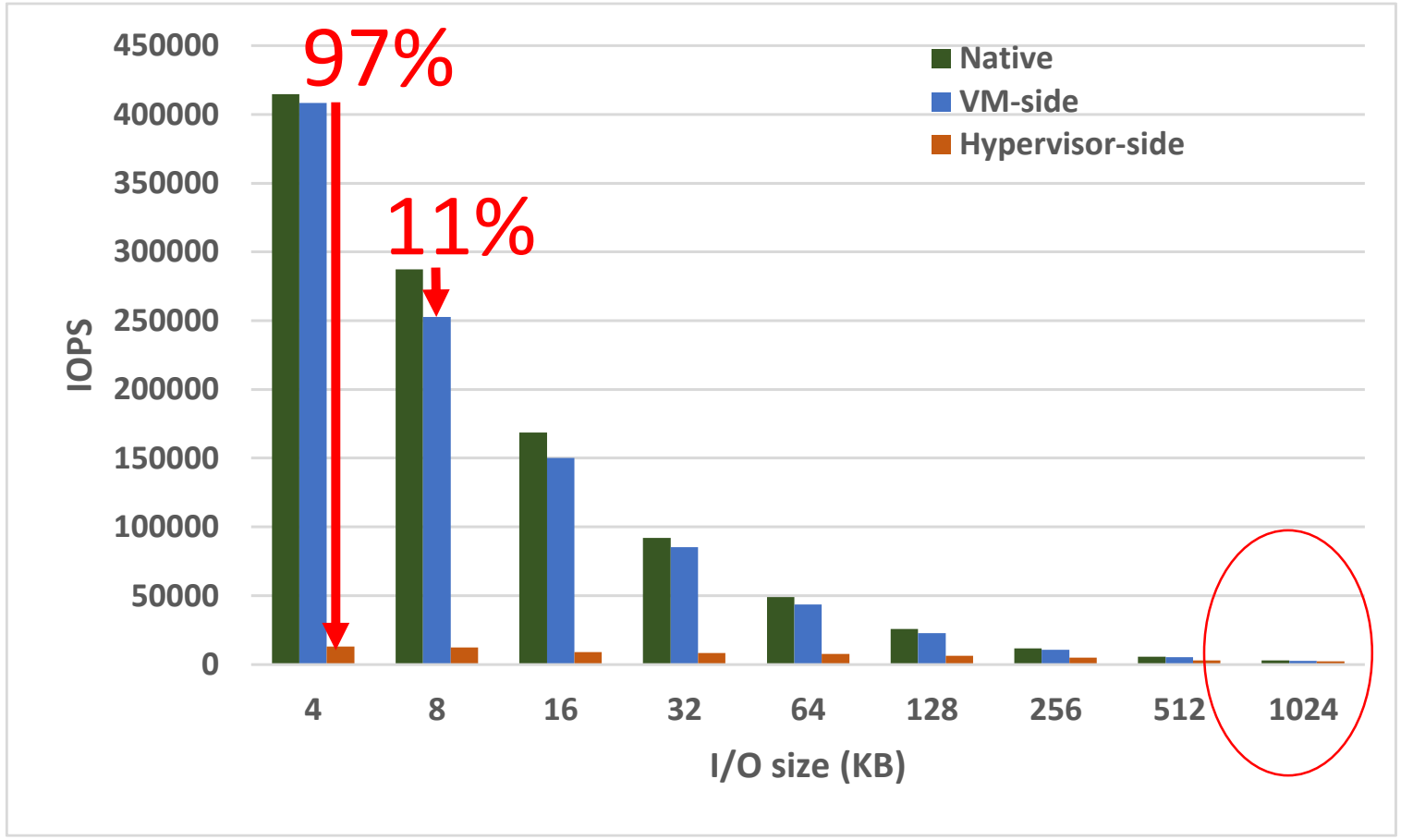
- Benchmark
  - *fio*

DRAM caches: VM-side vs. Hypervisor-side vs. Native

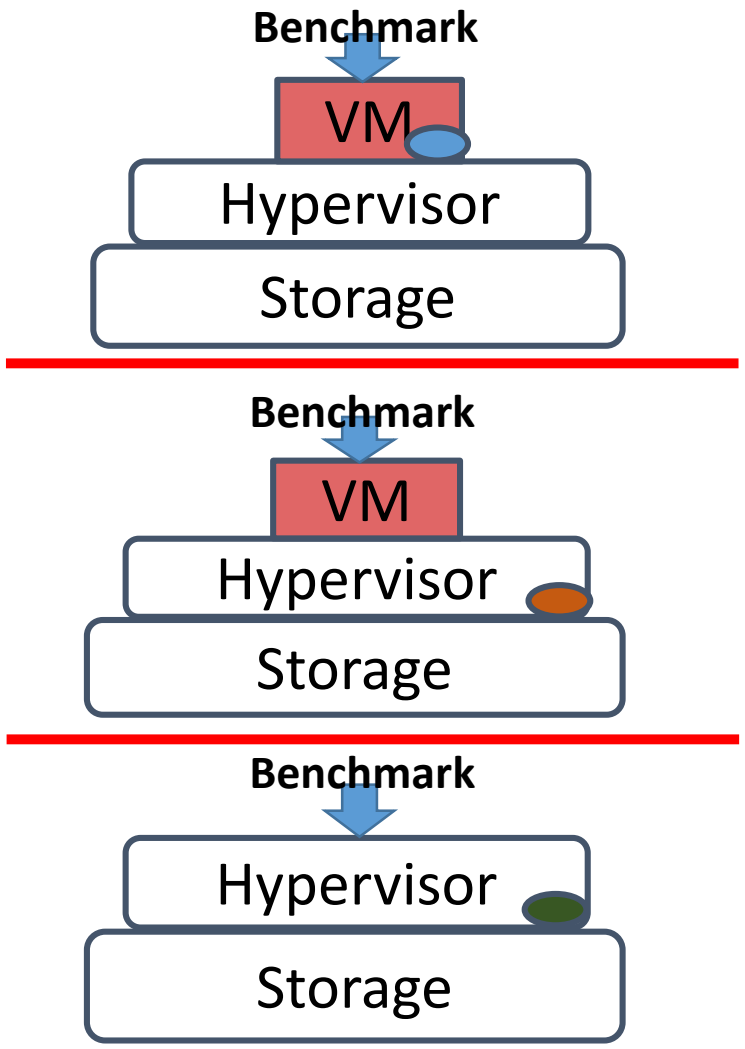


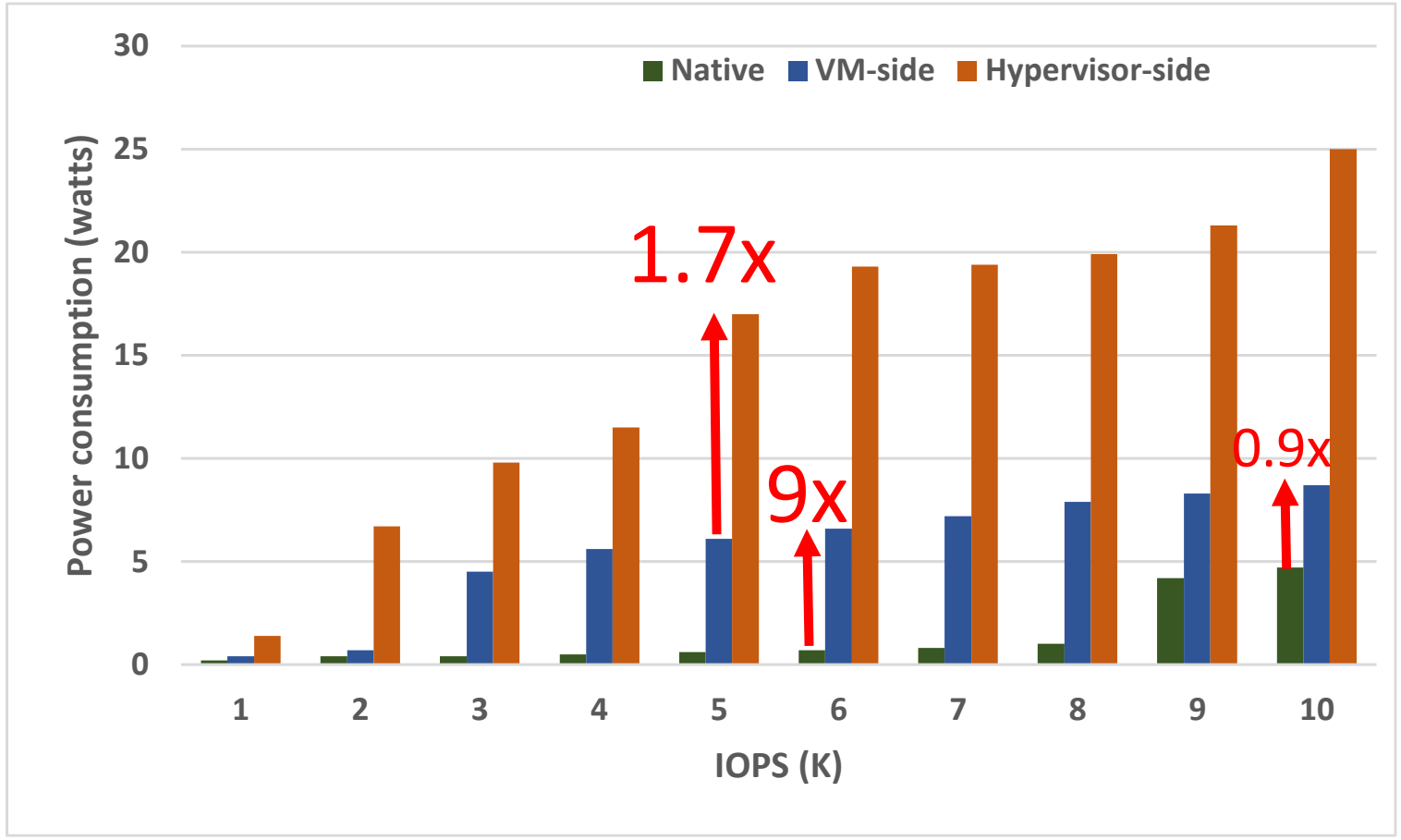
Performance of DRAM caches





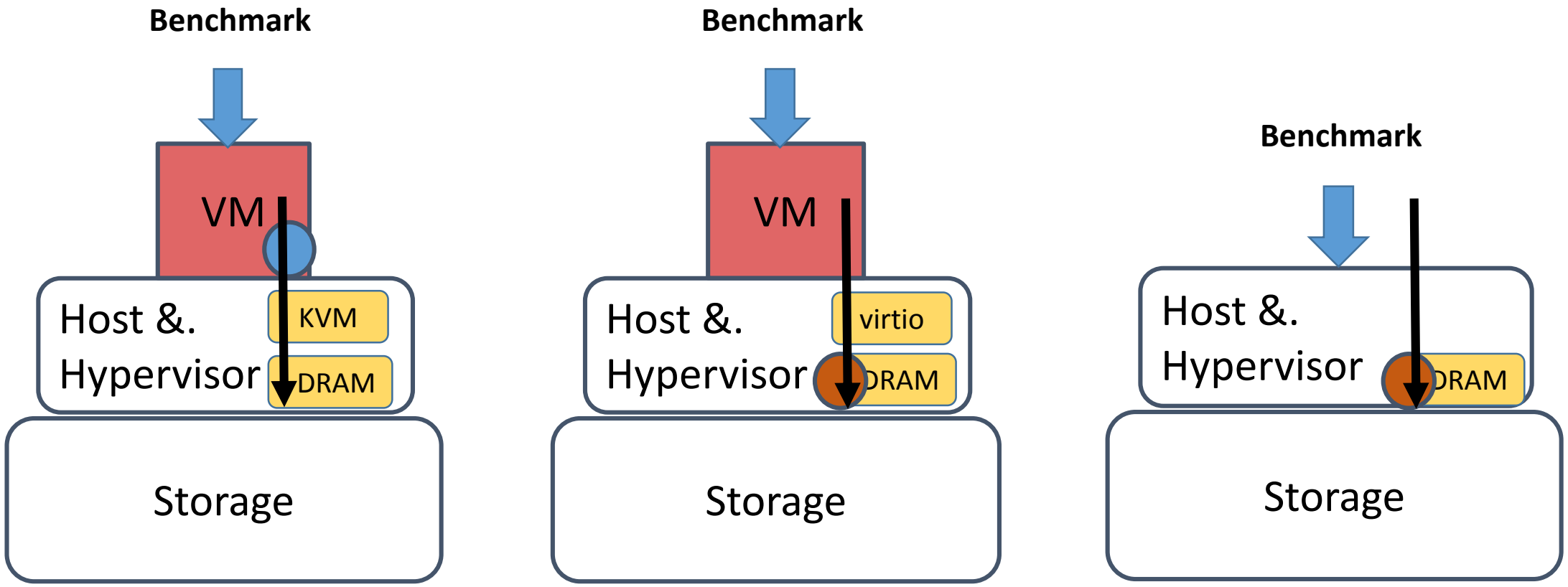
Power consumption of DRAM caches



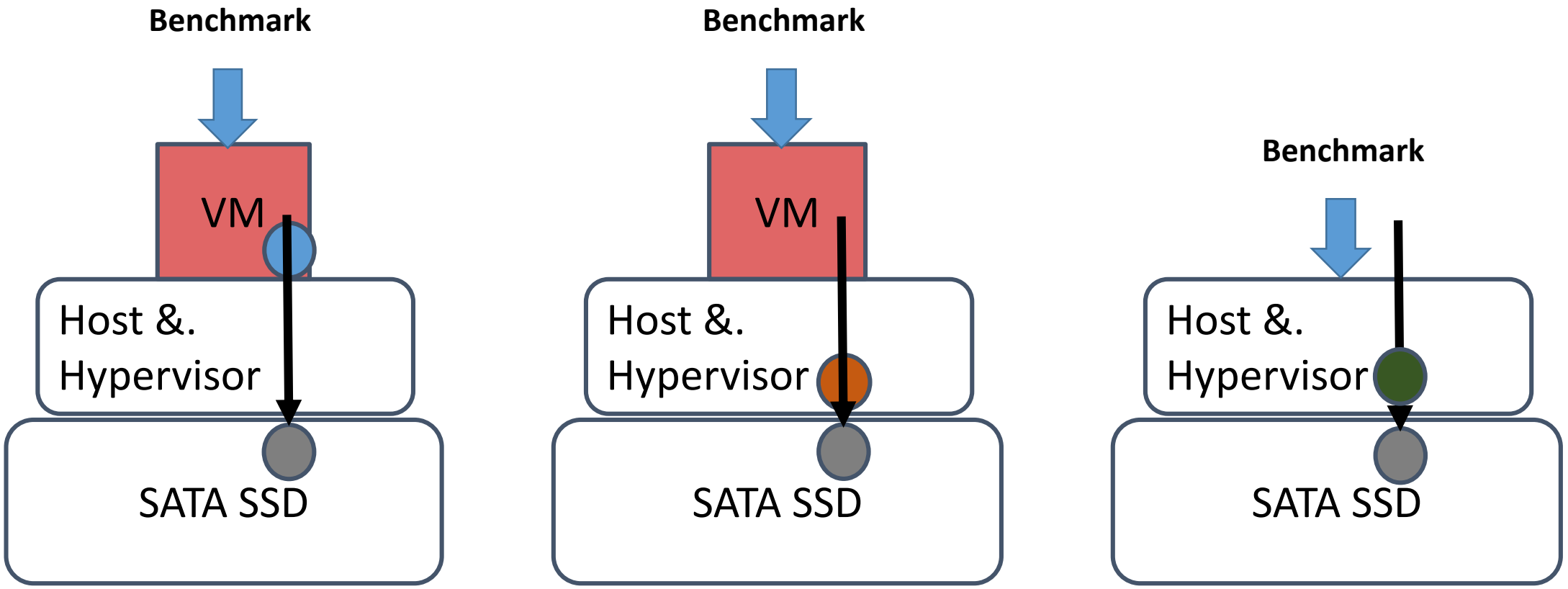


For DRAM caches, VM-side caching has better performance and energy efficiency than Hypervisor-side caching.

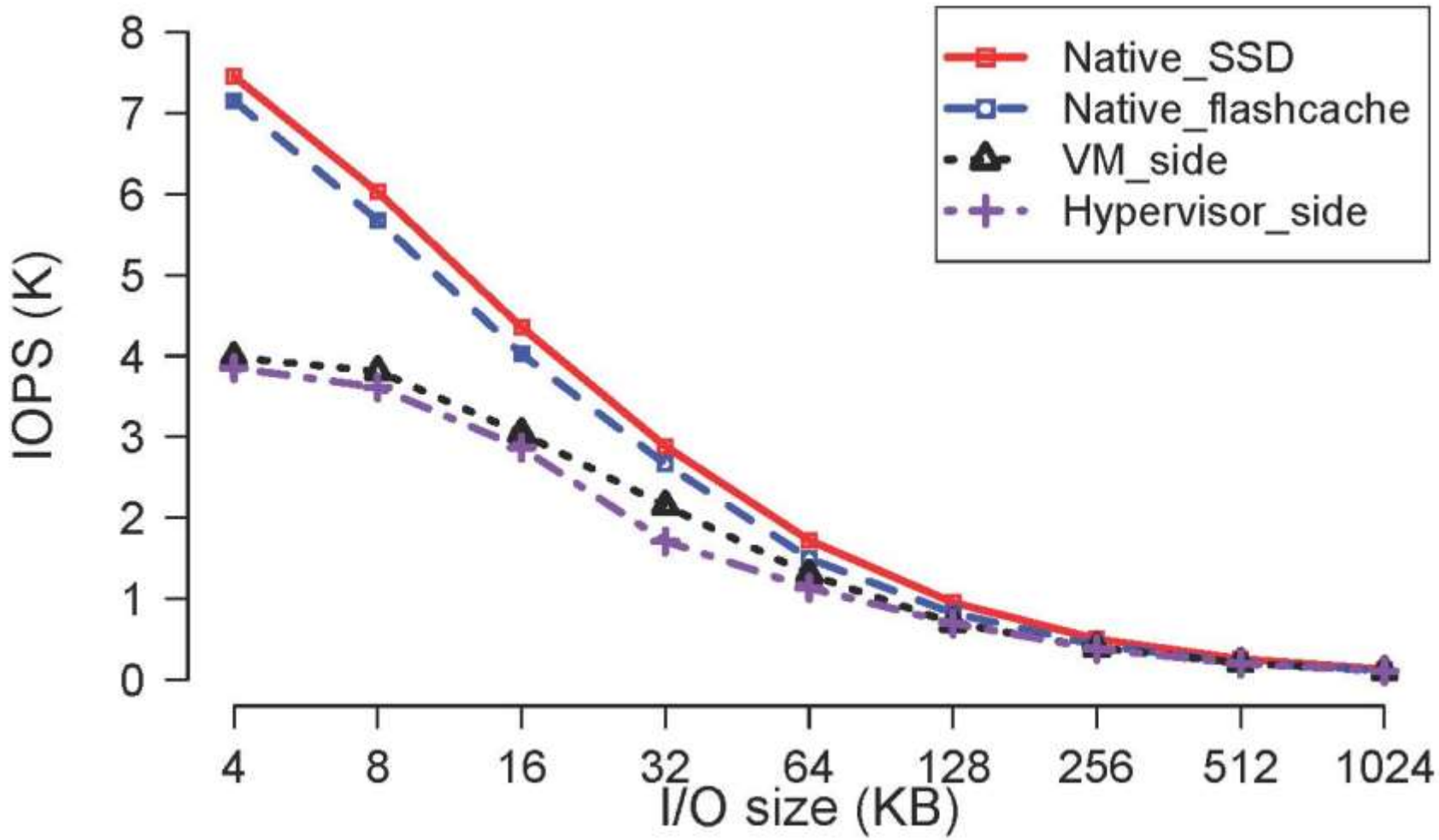
DRAM cache comparison: VM-side vs. Hypervisor-side vs. Native



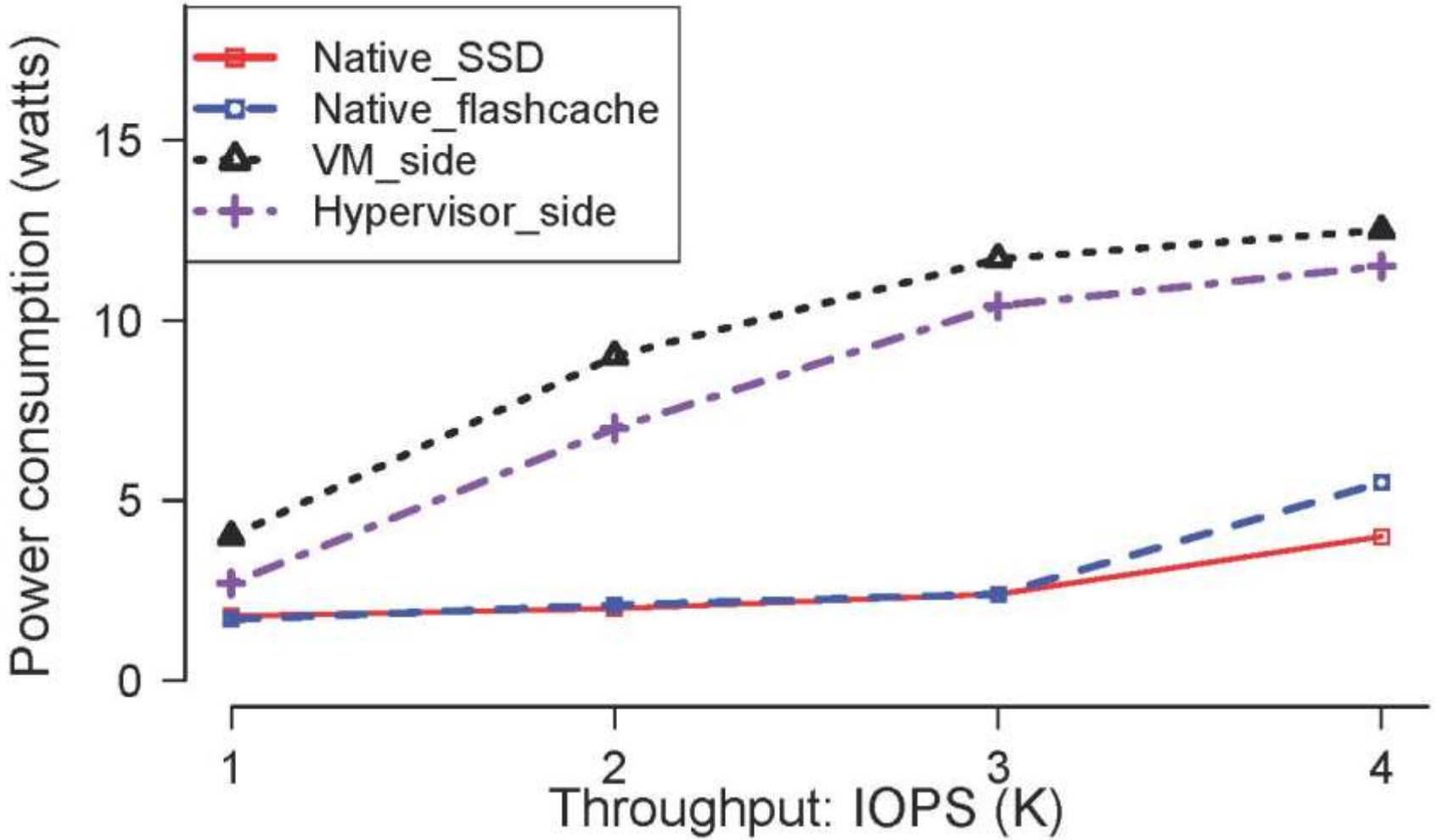
SATA SSD caches: VM-side vs. Hypervisor-side vs. Native



Performance of SATA SSD caches

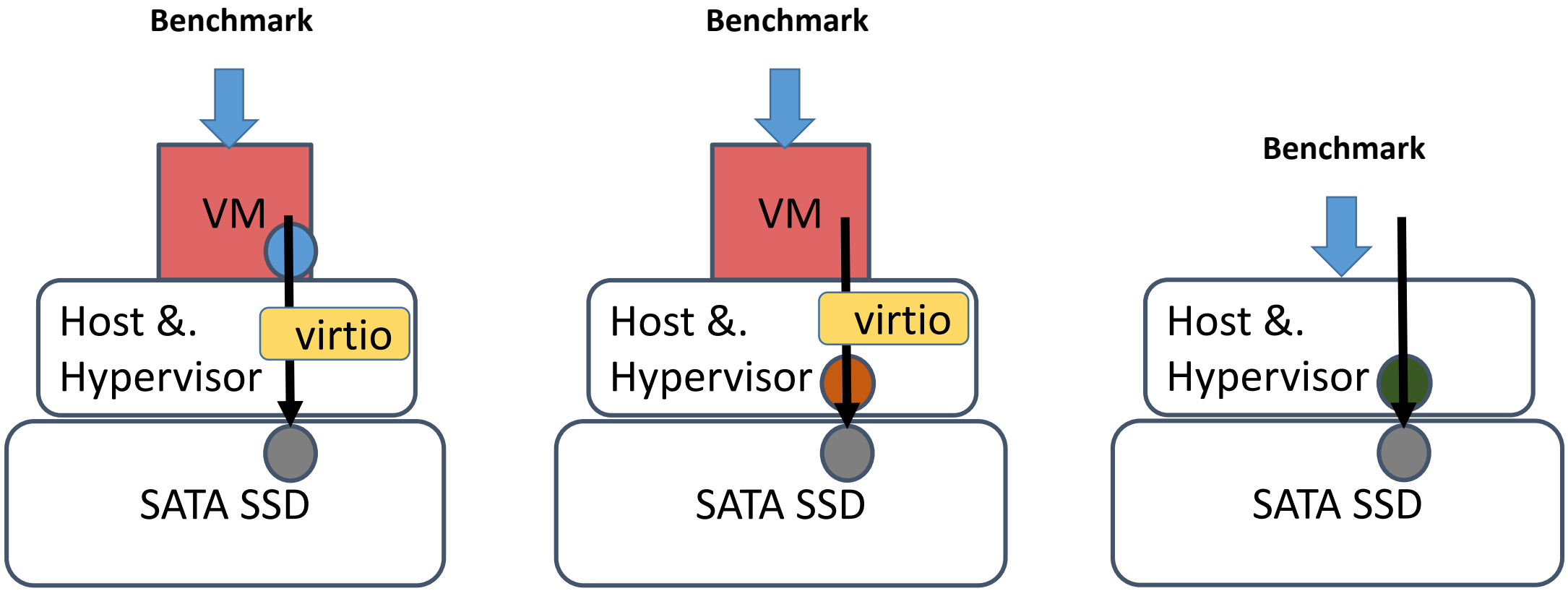


Power consumption of SATA SSD caches

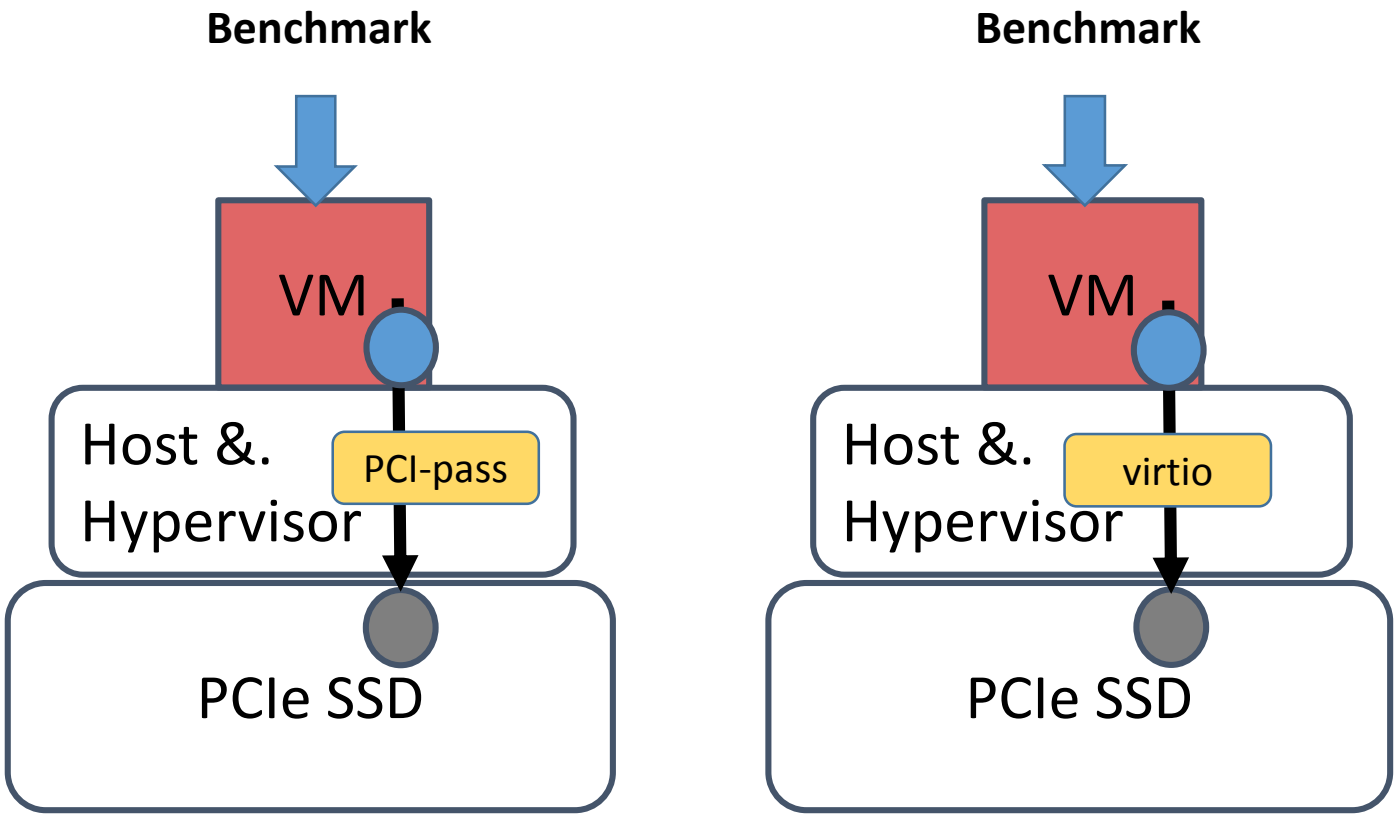


For SATA SSD based caches, VM-side caching and Hypervisor-side caching perform similarly.

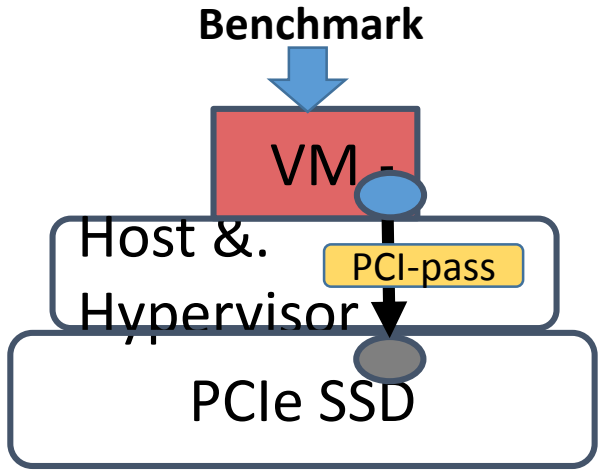
SATA SSD cache comparison: VM-side vs. Hypervisor-side vs. Native

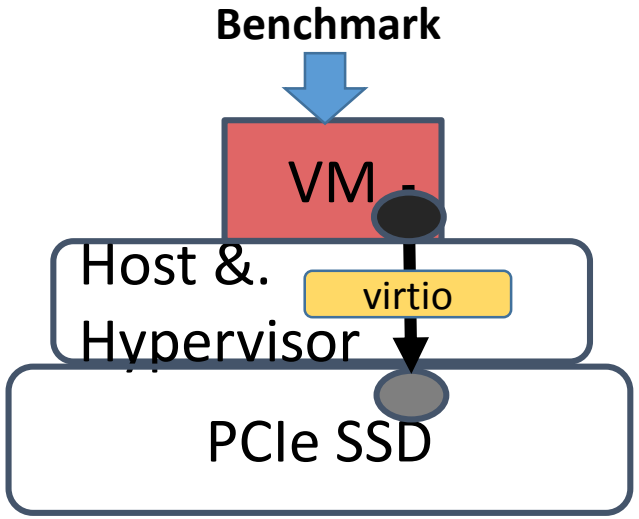


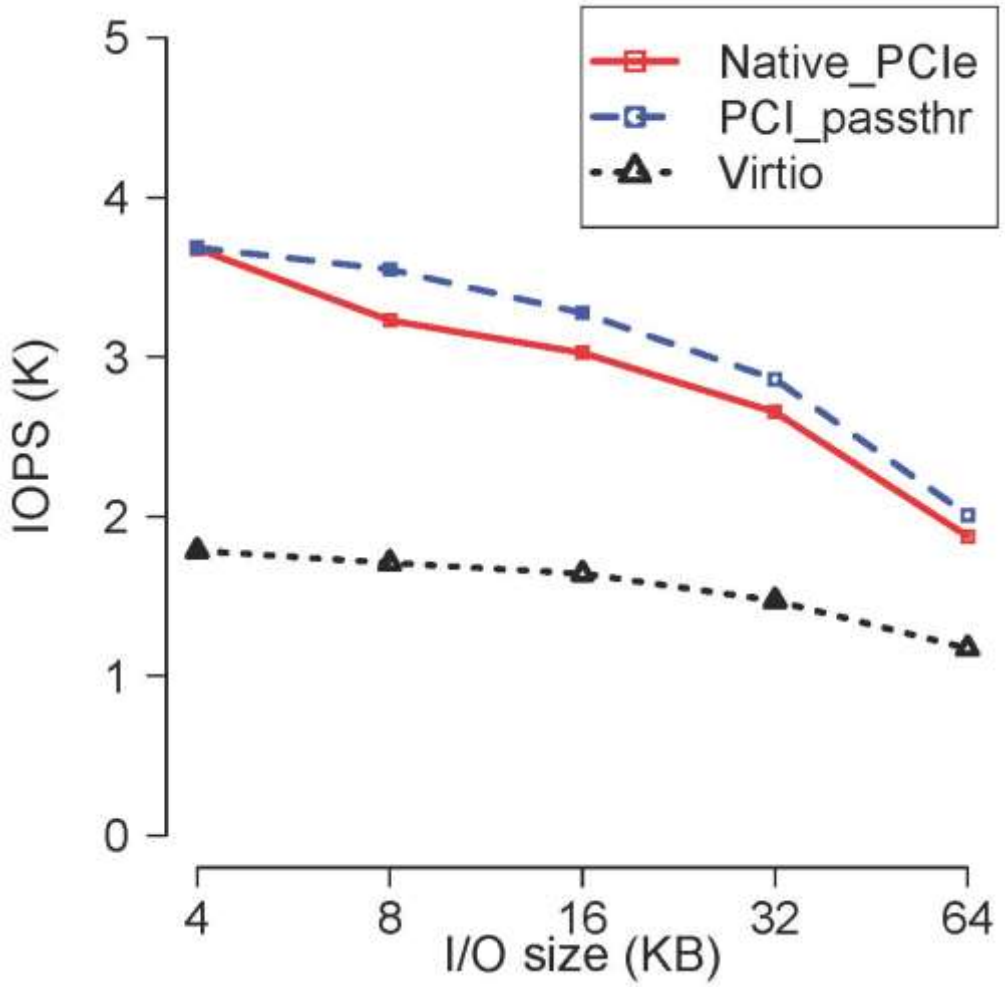
PCIe SSD caches: PCI passthrough vs. virtio



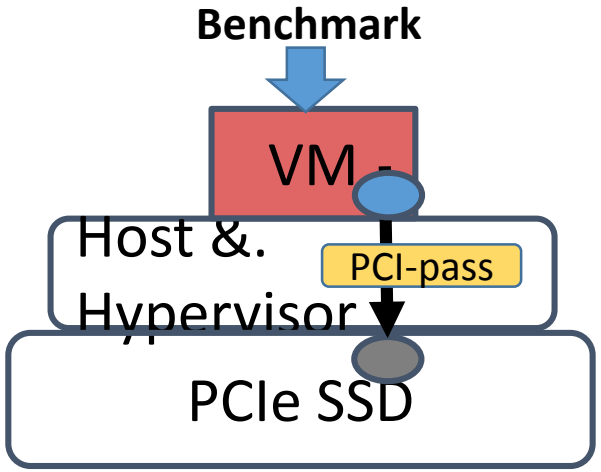
Performance of PCIe SSD caches (cont'd)

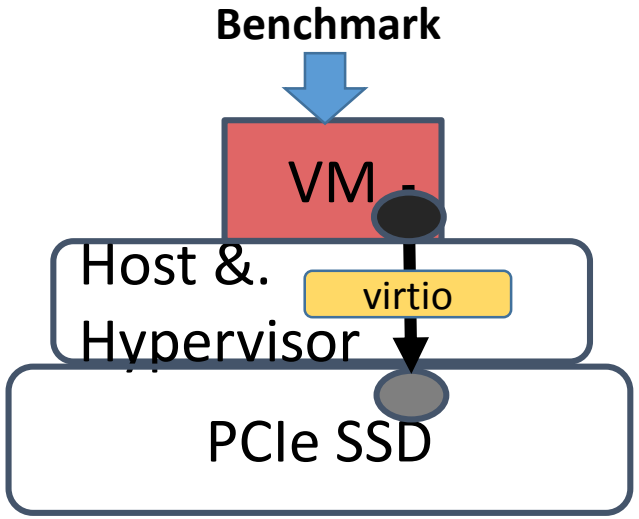


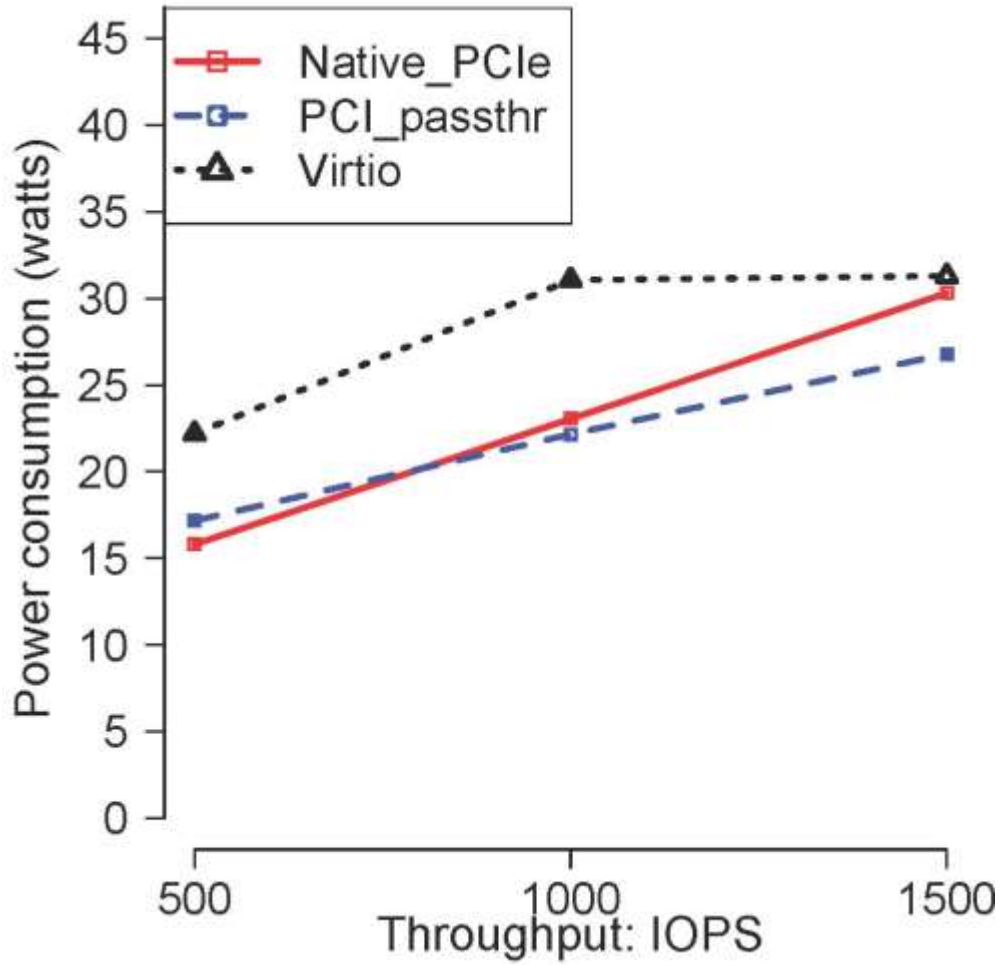




Power consump. of PCIe SSD caches (cont'd)







PCI passthrough enables the PCIe SSD allocated to a VM achieving near native performance.

Is hypervisor the best location to deploy caches?

SATA SSD	PCIe SSD	DRAM
Maybe YES	NO	Absolutely NO

What caps the performance of hypervisor-side caching?

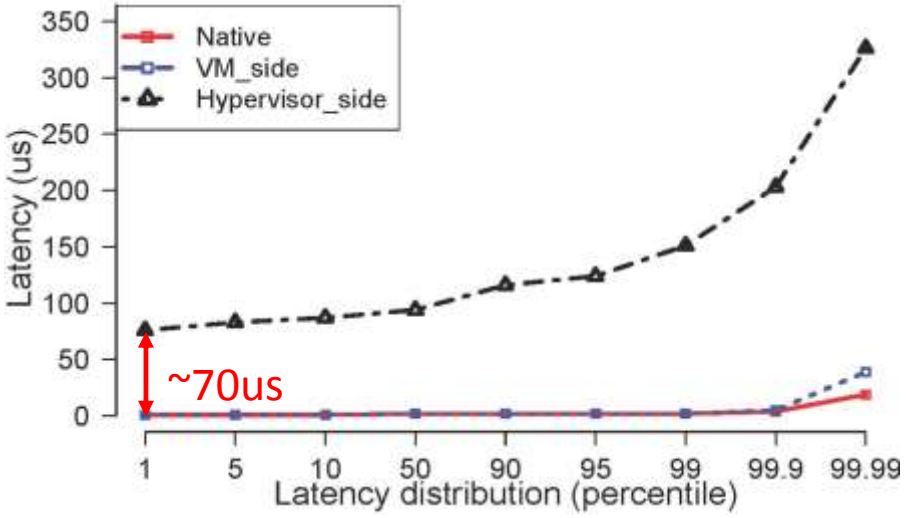
Virtual I/O path

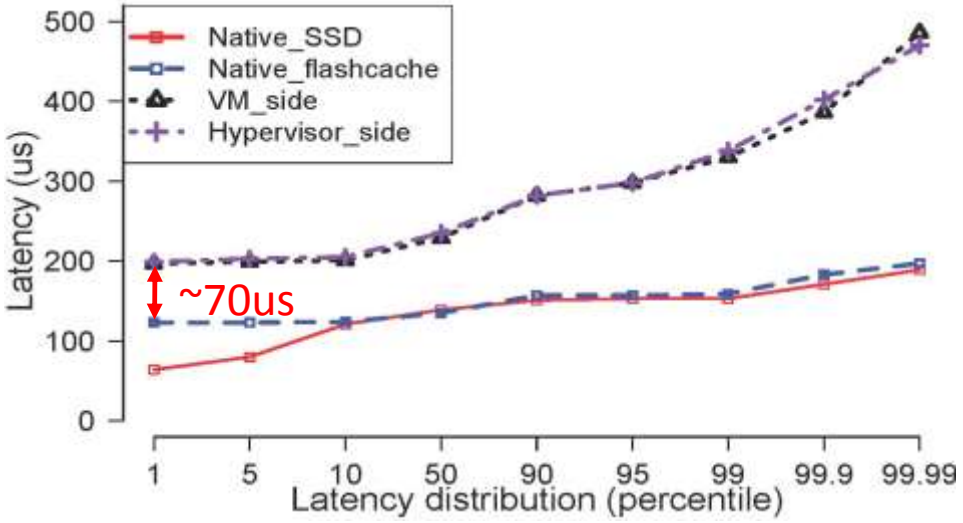
Virtual I/O path

- Challenge
  - ~60 us RTT (for *virtio*)

DRAM cache: 1-2 us

SSD: 40-100us





Virtual I/O path

- Challenge
  - ~60 us RTT (for *virtio*)
- Chance
  - Bandwidth is high

Overview

- Cloud storage stacks
- Impact of cache locations on system performance and energy efficiency
- Potential optimizations

Potential optimizations

- Resource Management
  - Allocating DRAM resources directly to VMs
- Guest OS Optimization
  - VM-side block device read-ahead
- Hardware Support/ New Devices
  - PCI Passthrough
- Host OS/ Hypervisor Optimizations
  - Reducing Virtual I/O Overheads

Thanks & Questions