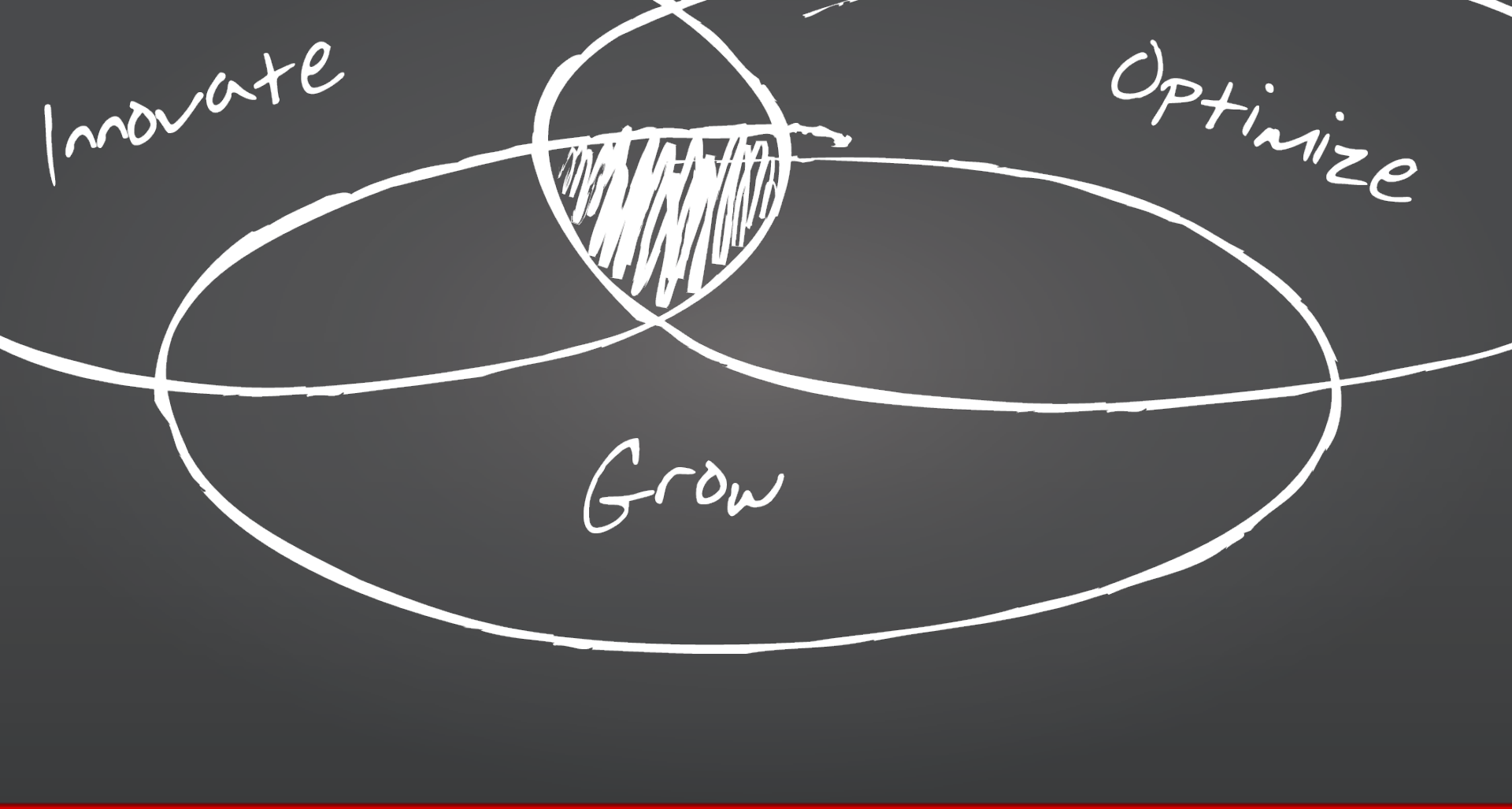


The total virtualization

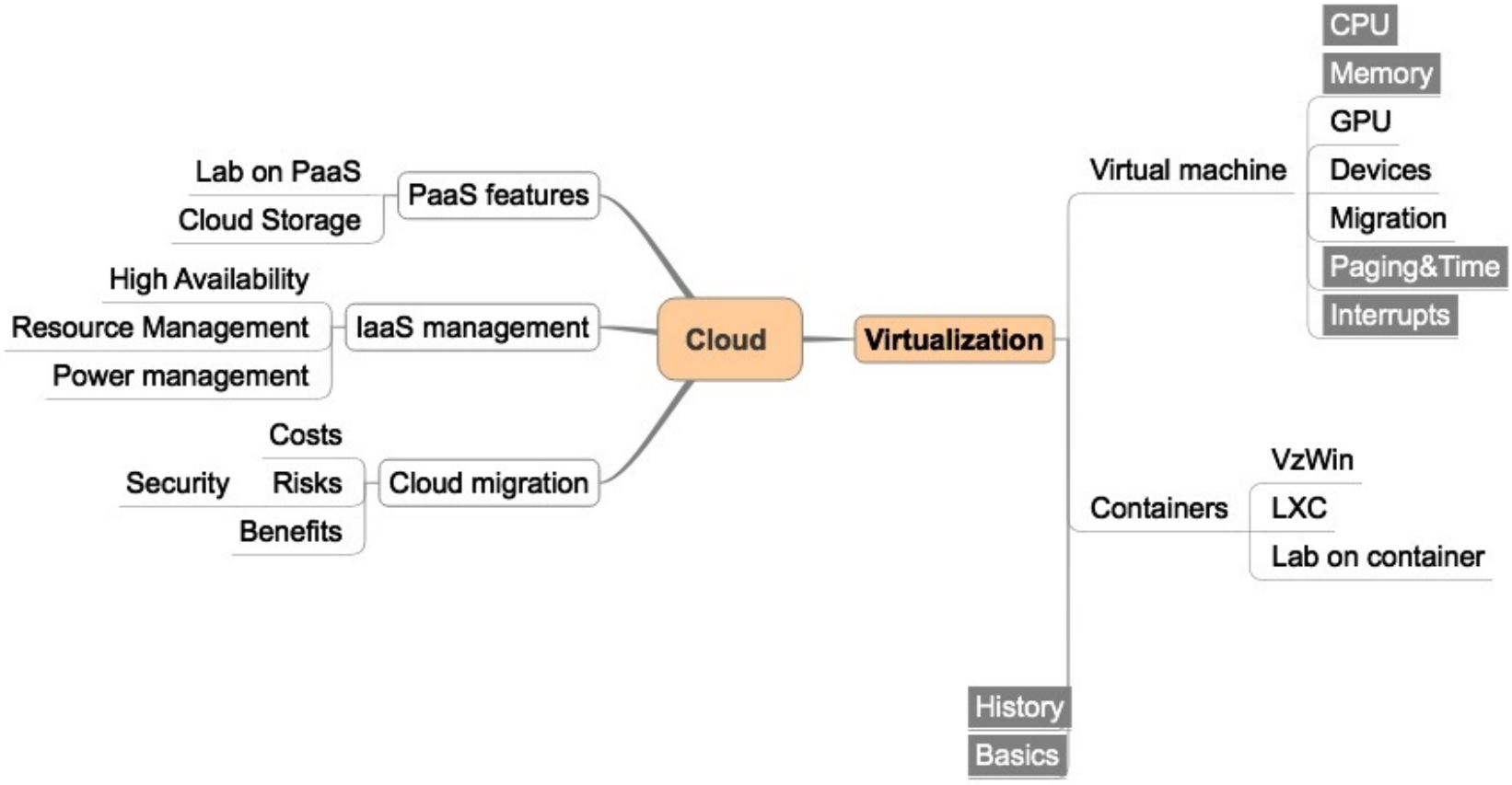
Virtualizing devices

Lecture 7

Content

- ✓ Virtualizing devices
- ✓ Virtualizing network
- ✓ Virtualizing HDD

Course overview



Effective working with network and HDD requires understanding of their low-level algorithms.

Basics of device interaction

Basics of device interaction

- ✓ Get device status
  - ✓ Polling
  - ✓ Interrupt-driven
- ✓ Set device cmd
  - ✓ I/O
  - ✓ Memory-mapped I/O
- ✓ Data exchange
  - ✓ I/O / Interrupt
  - ✓ DMA, Memory-mapped I/O

Device virtualization

- ✓ Deliver device interrupts
- ✓ Handle device I/O
- ✓ Handle device DMA, Memory-mapped I/O

Device virtualization

- ✓ Deliver device interrupts
  - ✓ In-time delivery, low latency
- ✓ Handle device I/O
  - ✓ Lots of code, meet specs
- ✓ Handle device DMA, Memory-mapped I/O
  - ✓ Lots of code, meet specs
  - ✓ Performance (speed-up data exchange between guest device and the real one)

Device virtualization: alternatives

- ✓ Device pass-through
- ✓ Device paravirtualization

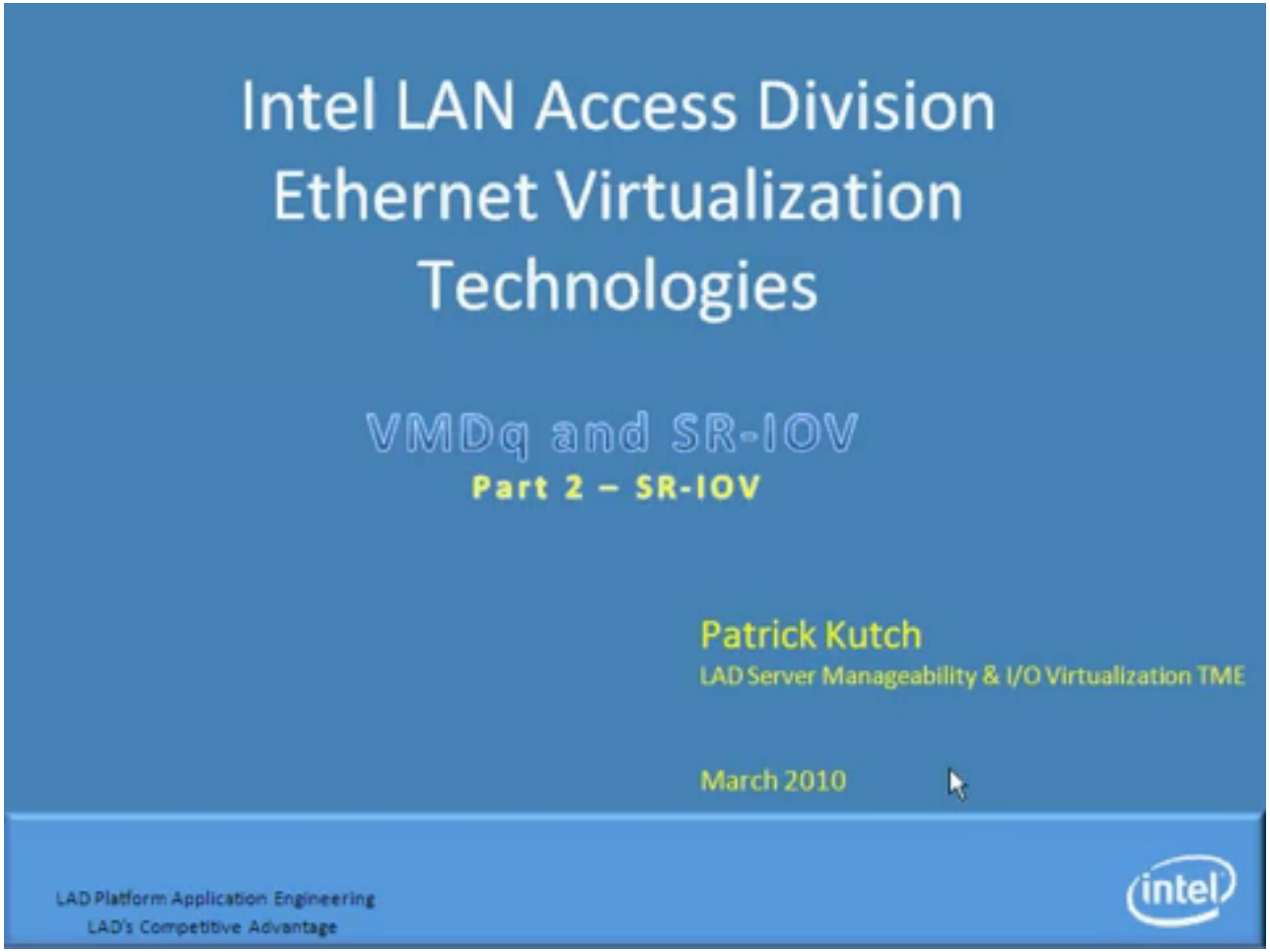
Paravirtualization approaches

- ✓ Hypercall on OS level
  - ✓ Xen
  - ✓ Hyper-V
- ✓ Paravirtualizing driver
  - ✓ toolgate work – vmware/parallels/vbox
- ✓ A special virtual device
  - ✓ Virtio (KVM)

Device pass-through (1)

- ✓ Steal the device from the host
- ✓ Share the host device with a special chipset features

Device pass-through (2)



Device pass-through: challenges

- Not all devices fit PCI standard
  - Reading configuratuion space not by means of OS API
  - Not all devices support reset function
- Not all devices could be taken from host OS
- Challenges with interrupt delivery (int remapping)

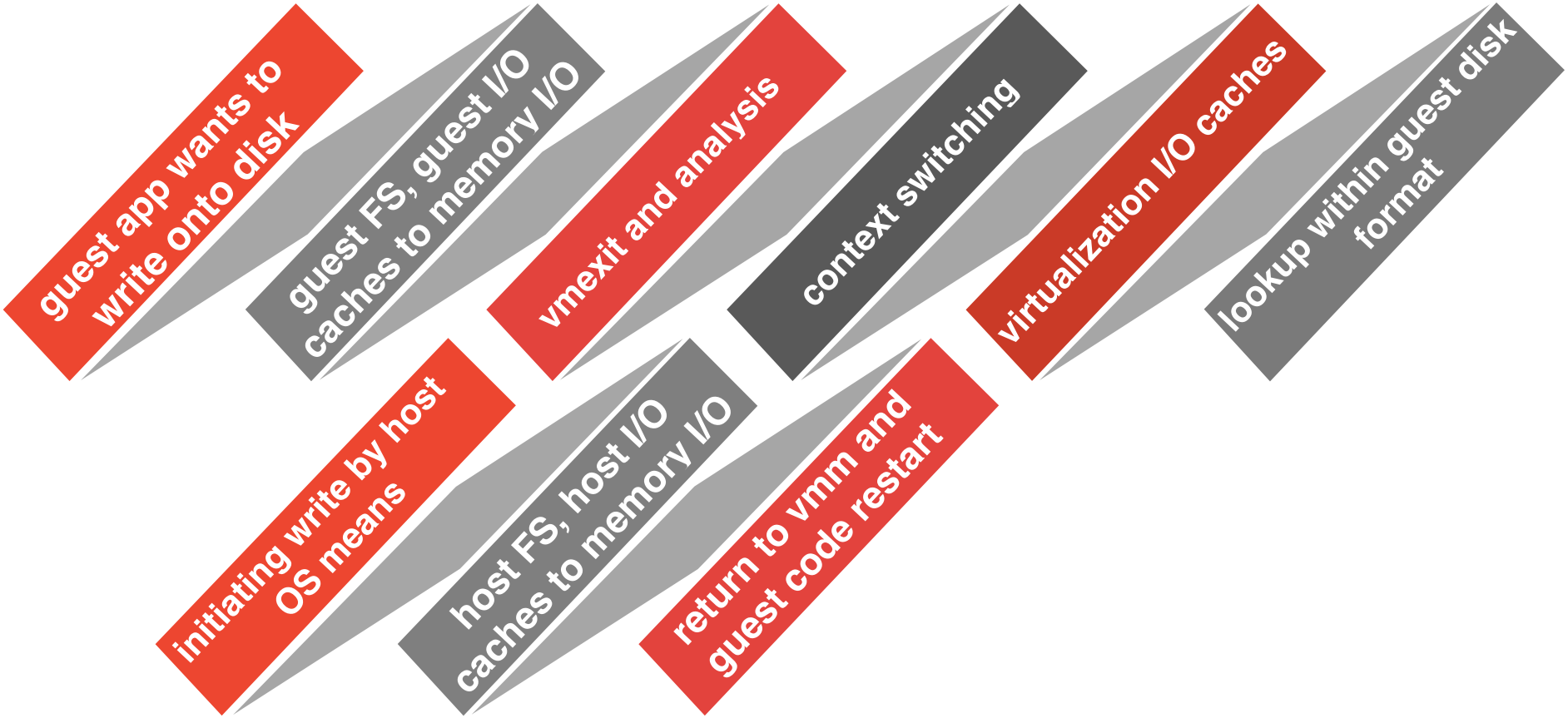
Hard disk drive

- Hard drive characteristics:
  - Time to access data
  - Seek time
  - Latency
  - Data transfer rate
- Interfaces:
  - IDE
  - SATA
  - SCSI
  - SAS

HDD: scheduling

- ✓ FIFO, NOOP
- ✓ LIFO
- ✓ SSTF (shortest search time first),
- ✓ SCAN (Elevator = as head goes)
- ✓ Random
- ✓ CFQ (Completely Fair Queuing)

How virtualized disk write looks like?



HDD: performance

- ✓ Lookahead reading
- ✓ Lazy write
- ✓ Packaging
- ✓ Paravirtualization
- ✓ Performance vs stability compromise

HDD: format expectations

- ✓ Performance
  - ✓ Merging snapshots
  - ✓ COW/COR for fast instancing
- ✓ Feature extendibility
- ✓ Resizing
- ✓ Optional compression & encryption
- ✓ Fast migration

Practical: how to speedup virtualized disk

decrease context switching number

Use guest disk stack speedup (virtio), use host disk paravirtualization (vhost)

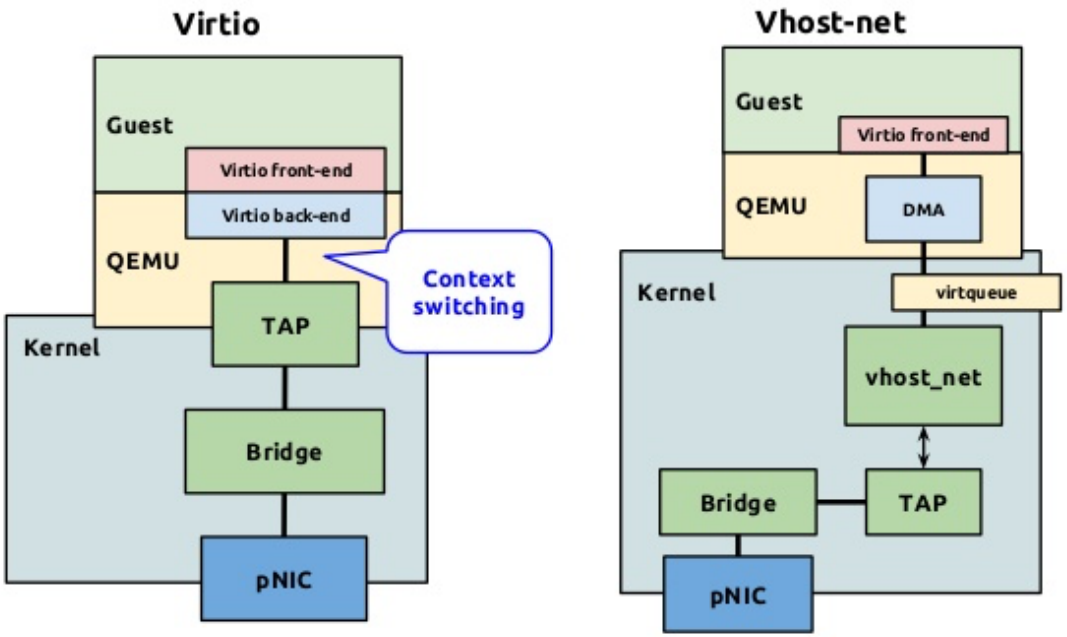
«fast» format of your HDD

turn off encryption, disable(?) linked clone, reduce snapshot tree

cache management

cache=None, aio=Native, metadata cache for qcow2

qemu+KVM epic



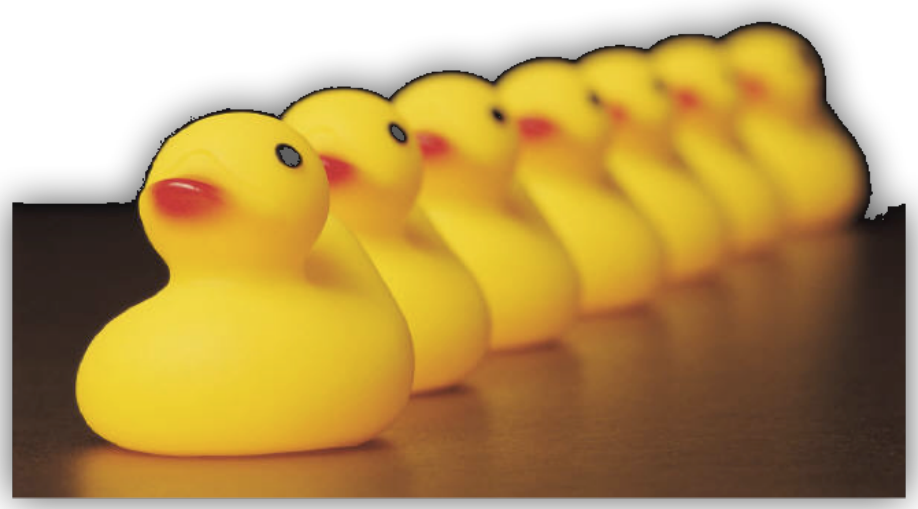
Conclusions

We all know that both disk and net are slow in comparison to memory and CPU. But each device and OS level try to compensate it by different algorithms. Aligning those algorithms with your usage is critical for the system performance

Questions?

Assignment 5 (virtualization)

- ✓ VirtIO: describe inflation and deflation process for virtIO balloon including guest Linux and host KVM part.



Projects

Project. Bufferbloat

Investigate bufferbloat problem. Present it to other students with a detailed comparison of existing solutions (CoDel & Co)