1. Virtual Machines

Virtual Machines

- A virtualization layer is interposed between the hardware and the operating systems
- Multiple operating systems can run on the same hardware simultaneously
- They can be the same o/s or different
- Each is isolated from the others and unaware of their existence
- A Virtual Machine Monitor is needed to accomplish this: The VMM or Hypervisor
- The overhead must be reasonably small – this has driven changes to chip design to support virtualization

(Intel VT (codenamed Vanderpool) and AMD's is referred to as AMD-V (codenamed Pacifica)

The idea has caught on

- An old idea in fact IBM 370 in 1972! Xen is far more flexible
- Sun’s VirtualBox
- Vmware ESX Server
- Microsoft just released Hyper-V
- Xen is the most widely used by far – available as open source but now owned by Citrix Inc.
- Developed at Cambridge University

Why only use 1 machine?

- Hypervisors can cooperate across a cluster or farm of servers

- Hypervisors can move a virtual environment from machine to machine

- This can be done in say 200mS – the user doe s not notice a delay!

- Why is this a brilliant idea?

Scalability and Robustness

- One can load balance the machines
- One can add more machines to the cluster easily offering more performance
- Trees of clusters are possible for really high performance
- A busy machine can offload some processes to a less busy one
- An unreliable machine or one suffering power failure can migrate its processes

2. Introduction of Hypervisor

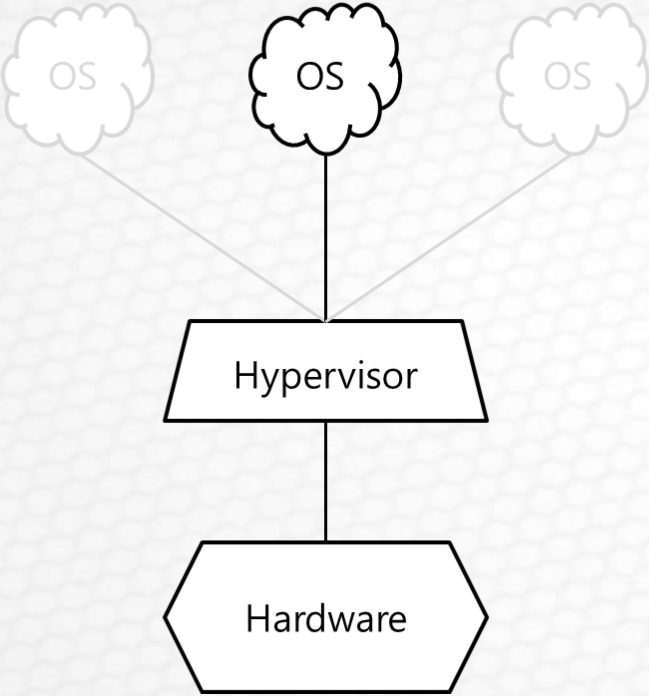
What is Hypervisor?

- H/W virtualization techniques allowing OS termed guests, to run concurrently on a host computer

- Type 1 Hypervisor (Native or Bare Metal)
  - Run directly on the host's H/W to control the H/W and manage guest OS
  - Citrix XenServer, VMware ESX/ESXi, Microsoft Hyper-V

- Type 2 Hypervisor (Hosted)
  - Hypervisor run within a conventional OS environment
  - Hypervisor level as a distinct second S/W level, guest OS run at third level above the H/W
  - KVM, VirtualBox

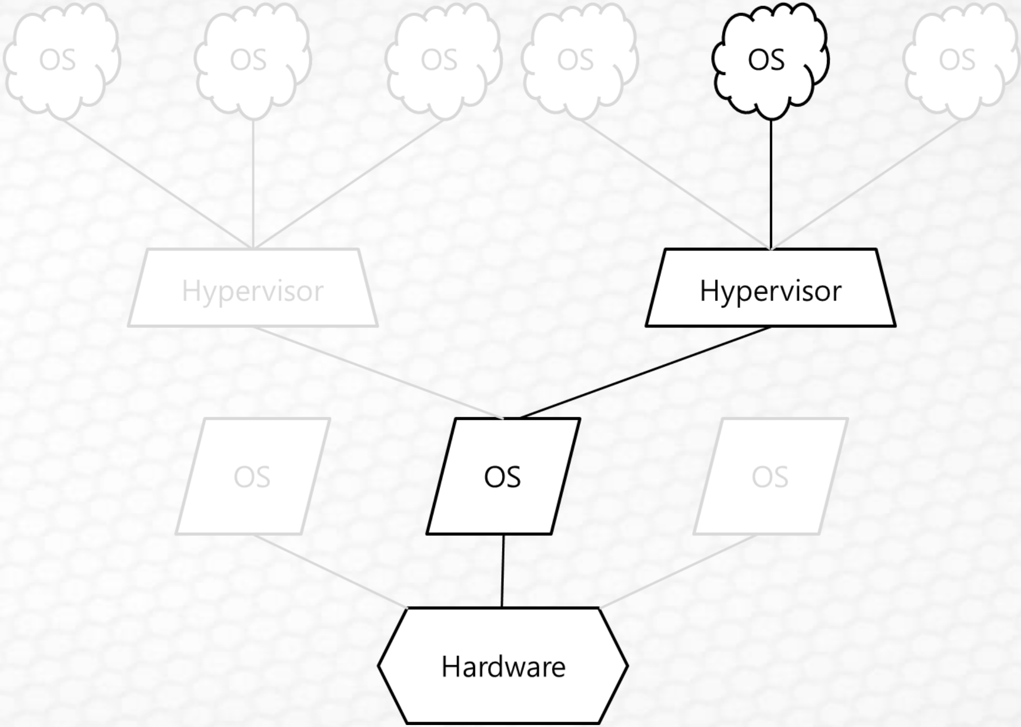
Types of Hypervisor



Type1

native

(bare-metal)

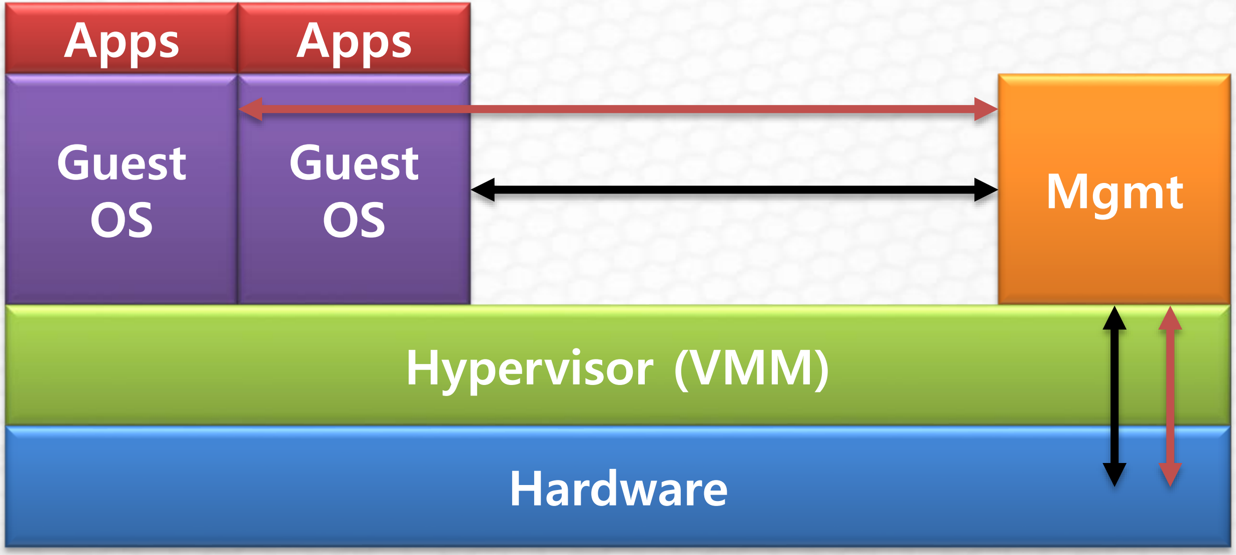


Type2

hosted

Full virtualization

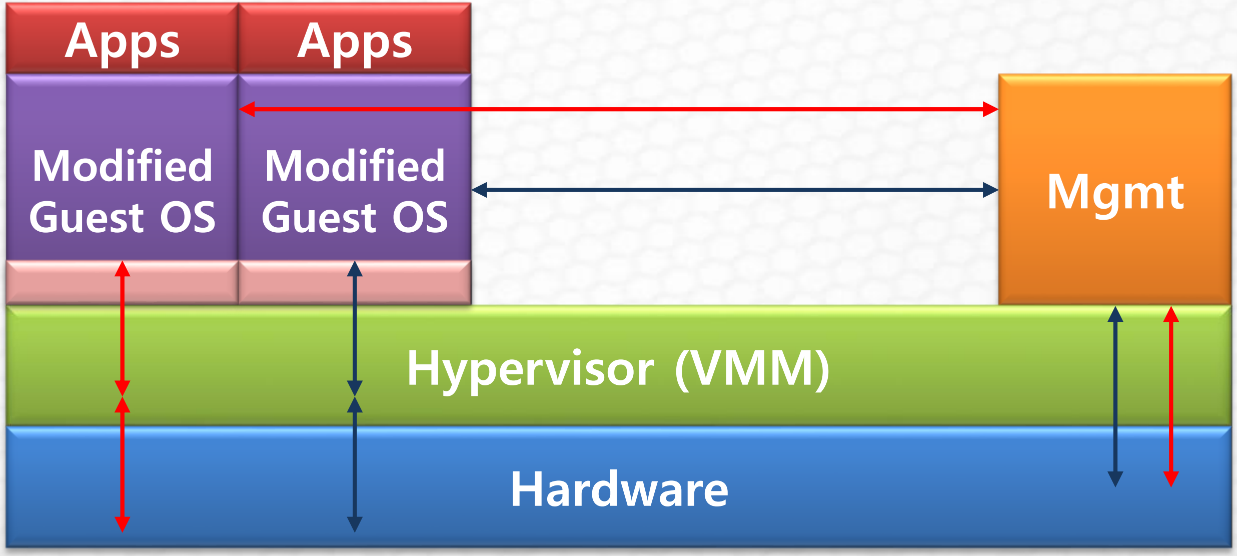
- Uses VM to *mediate* between Guest OS and H/W
- Fully virtualizes H/W, can support any type of OS with no configuration
- Certain Instruction Sets must be handled and trapped by hypervisor
- Low performance



Paravirtualization

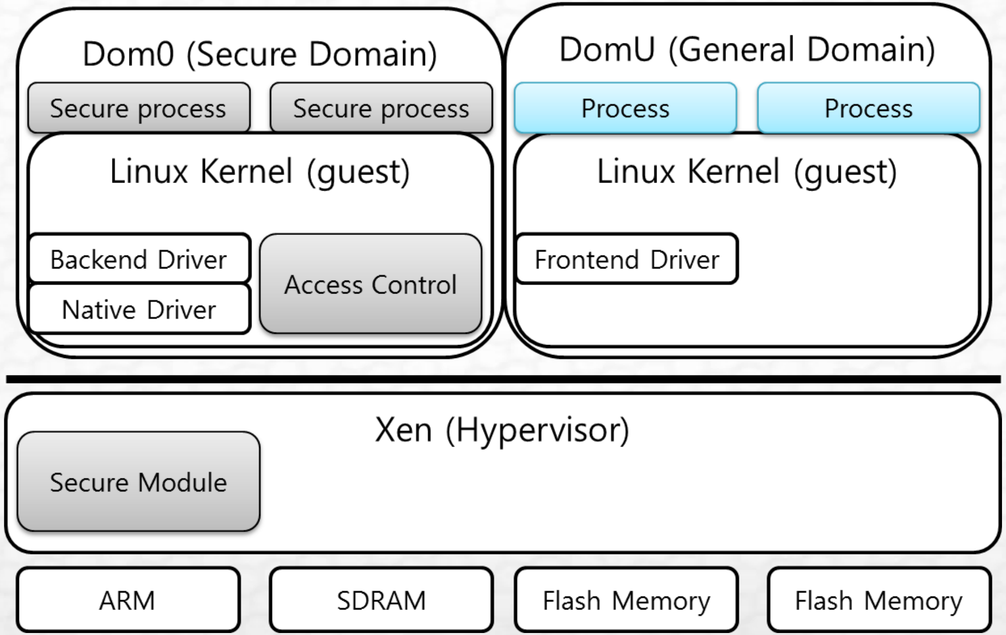
- Guest OS interacts with Hypervisor directly using *Hypercall*
- OS must be reconfigured with the corresponding Hypervisor
- Provides near native performance:

  Full virtualization < Paravirtualization



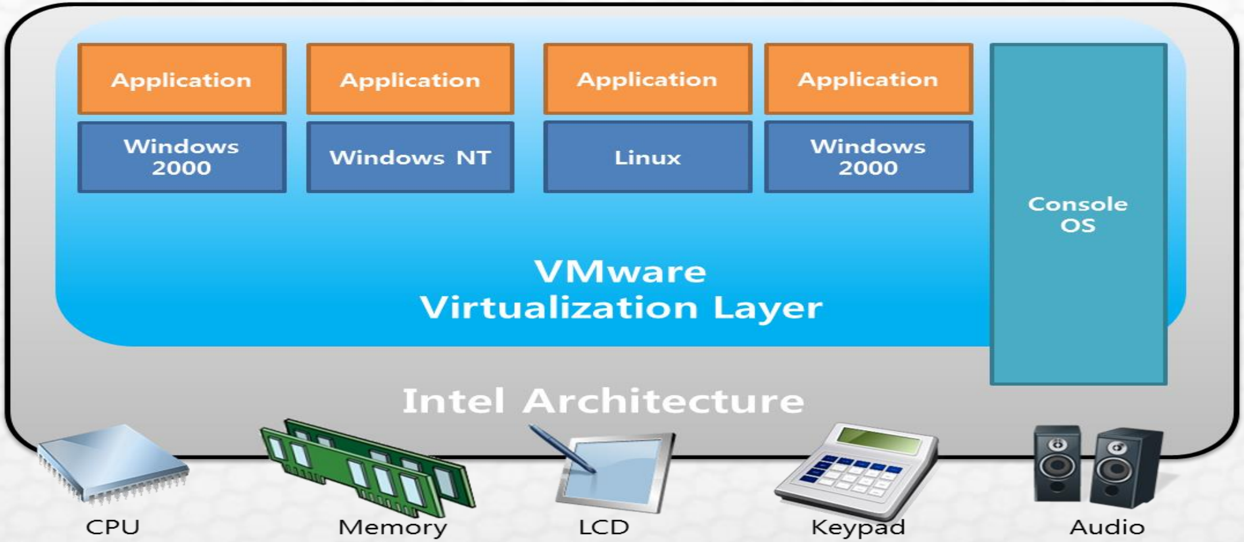
Xen Hypervisor

- Paravirtualization
- Dom0 is a privileged domain that can touch all hardware in the system
- Supports x86, x86-64, ia64 and PPC in varying degrees of maturity
- Supports live migration VMs



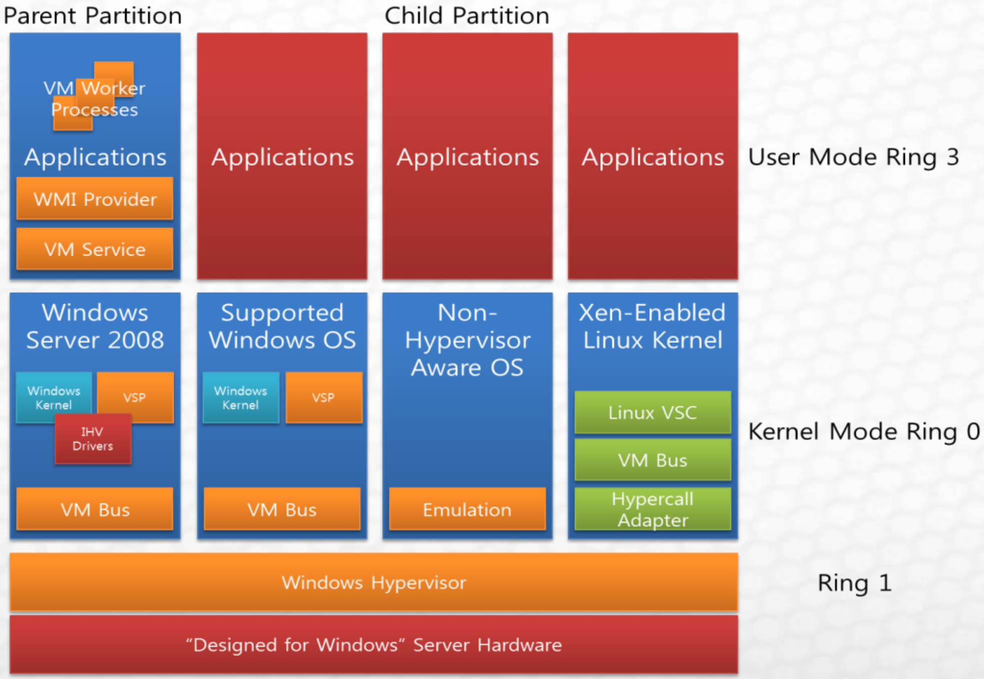
VMware ESXi

- Full Virtualization
- Relies on Linux OS called service console(Console OS) to perform some management functions including executing scripts and installing third-party agents for hardware monitoring, backup or systems management



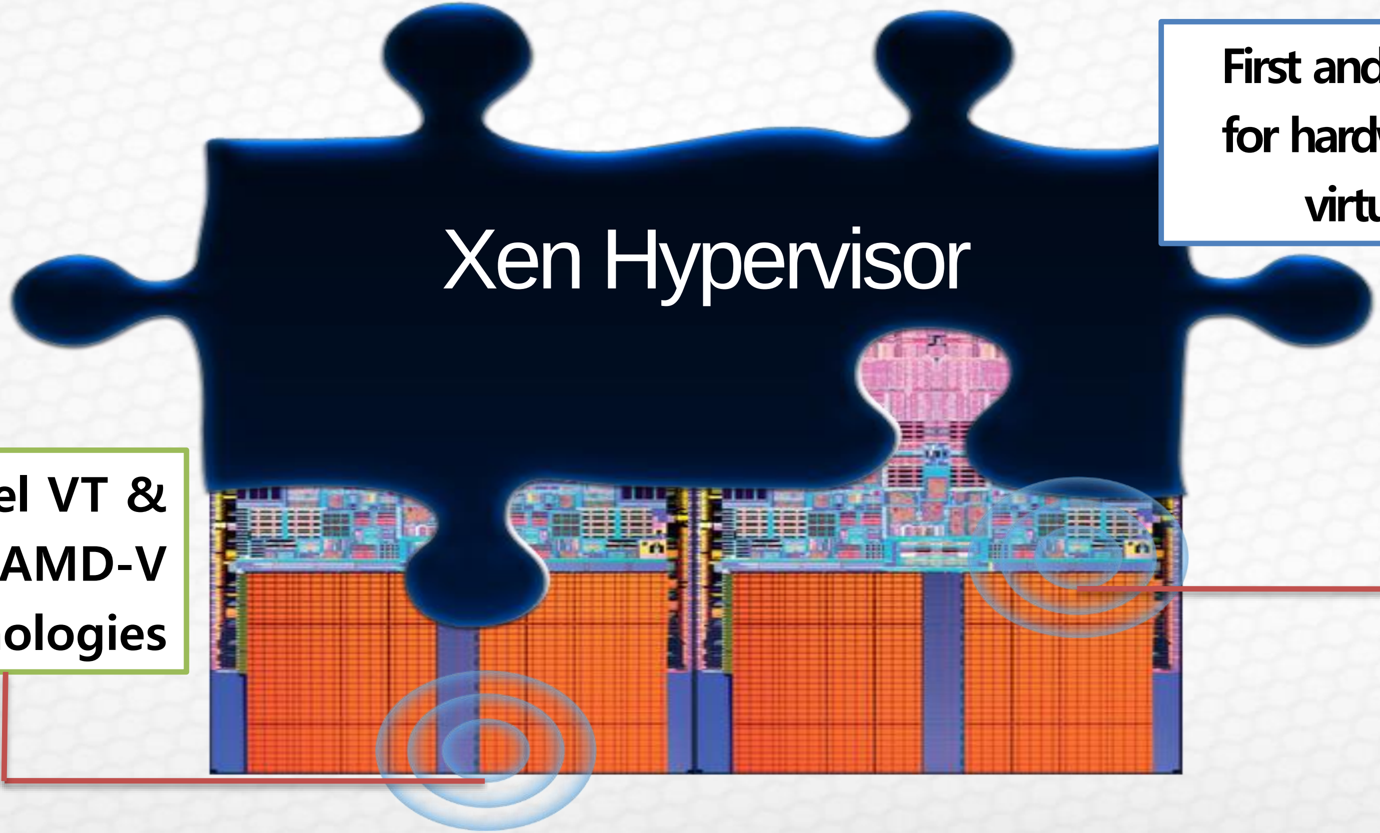
Hyper-V

- **Full Virtualization**
- **Stand-alone product by Windows Server 2008 R2**
- **Supports Live Migration**



3. Xen Hypervisor

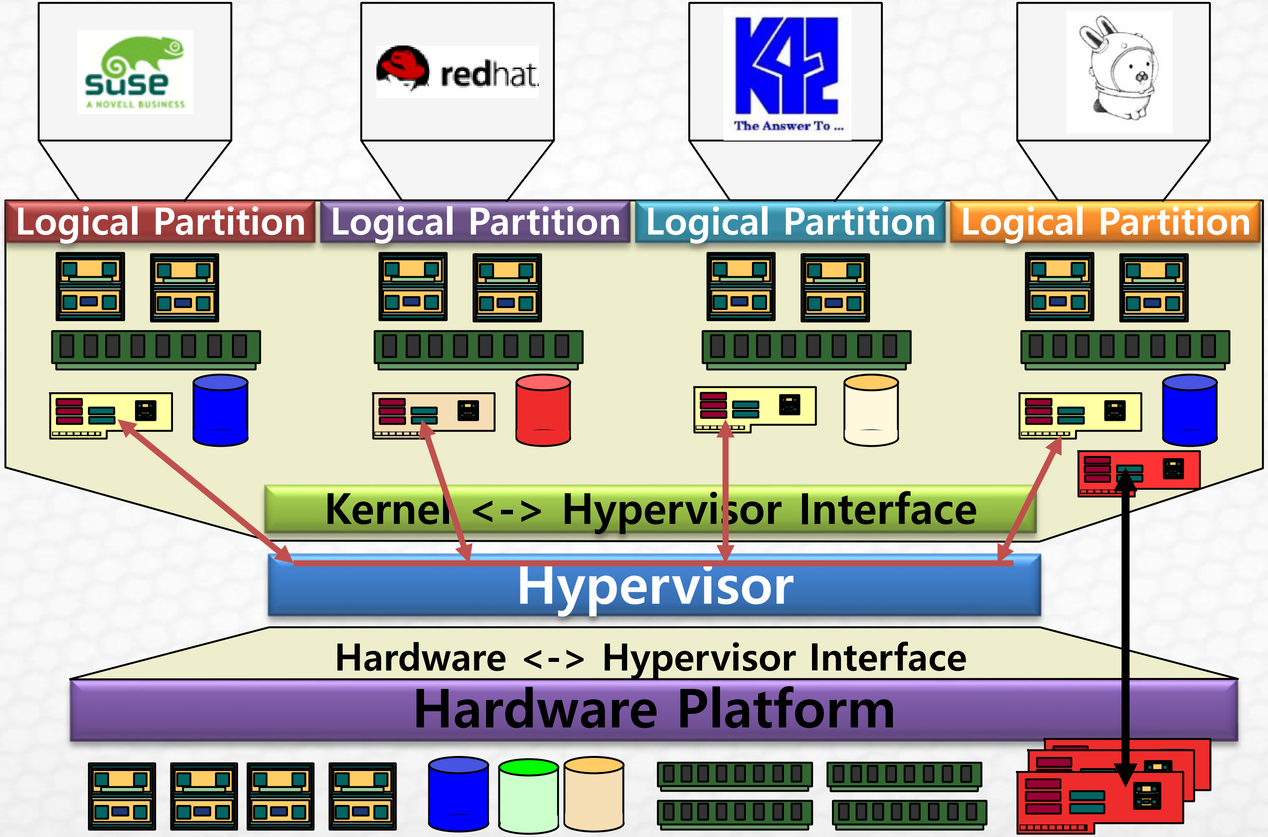
Xen Hypervisor



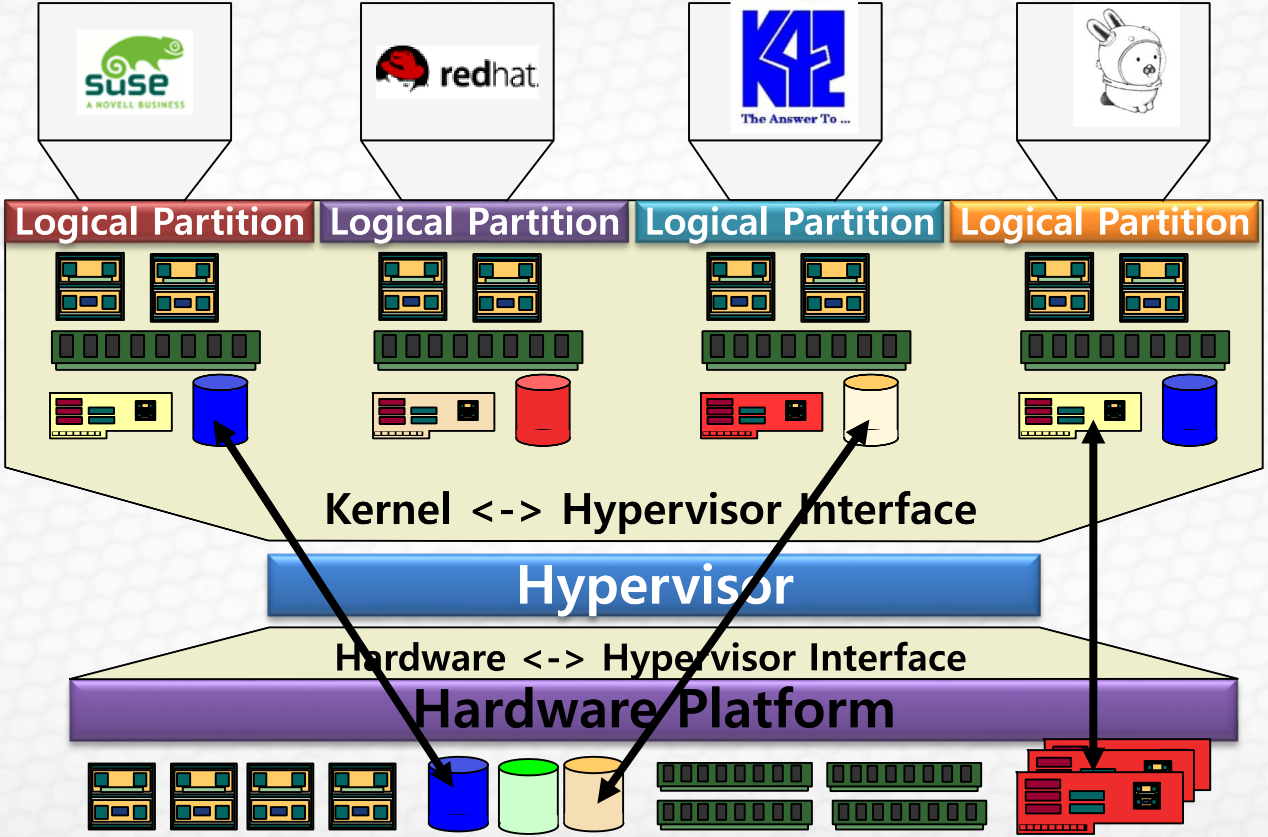
First and best support for hardware assisted virtualization

Intel VT & AMD-V Technologies

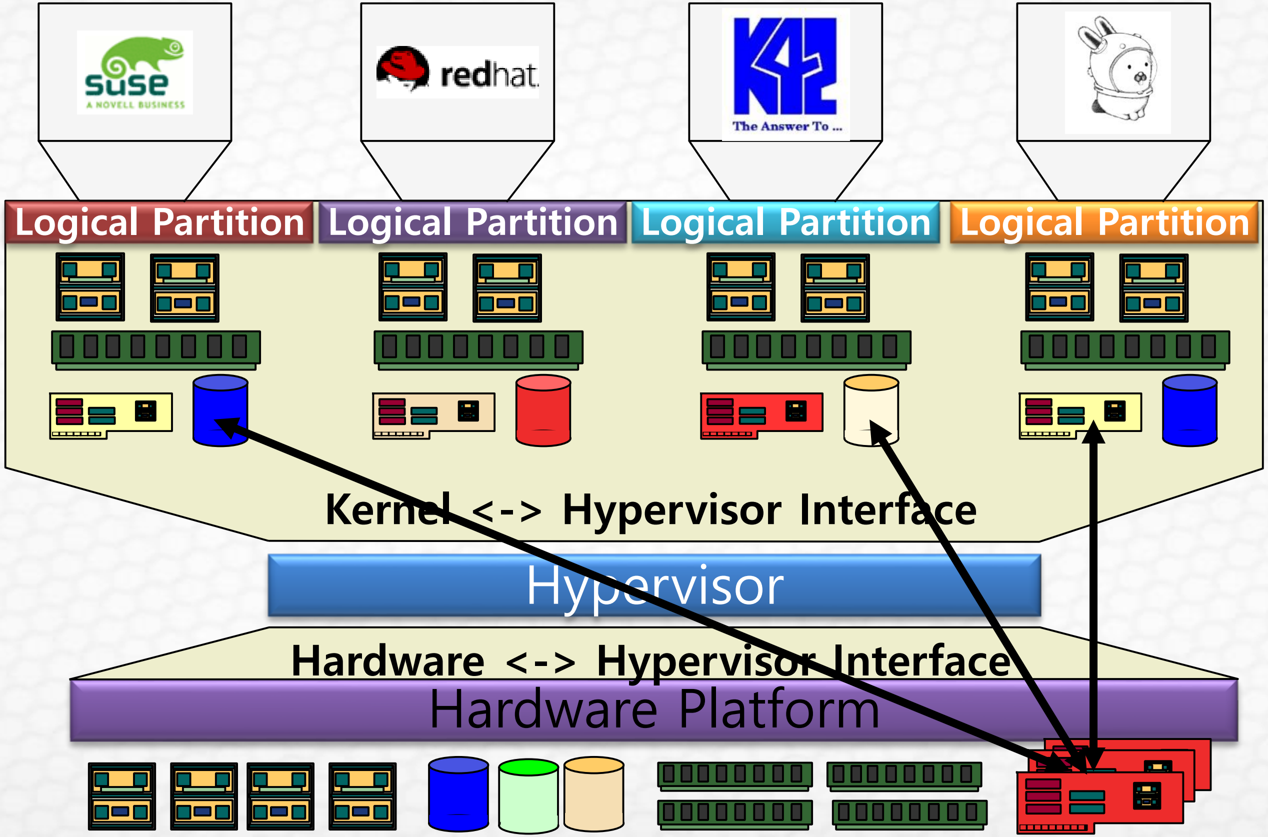
IO hosting


suse
A NOVELL BUSINESS
redhat
K42
The Answer To ...
Logical Partition
Logical Partition
Logical Partition
Logical Partition
Kernel <-> Hypervisor Interface
Hypervisor
Hardware <-> Hypervisor Interface
Hardware Platform

Device Partitioning with IOMMU



Self virtualizing devices



What we need to do

- Xen is the obvious choice.
- We need to help drive definition of hypervisor before it becomes too mature.
  - Investigate costs of their design decisions, and fix
- Need to drive definition of I/O virtualization and self-virtualizing devices.
- Determine set features we can use in common, e.g.:
  - One implementations of
  - checkpoint/restart/migration,…
    Gang scheduling of partitions.
- Hypervisor as a base is close to ready, making it first class platform for HEC will take investments..