Impact of Advanced Virtualization Technologies on Grid Computing Centers

International Symposium on Grid Computing 2009

Stefan Freitag

Robotics Research Institute
Dortmund University of Technology

23. April 2009

technische universität dortmund

Overview

Introduction

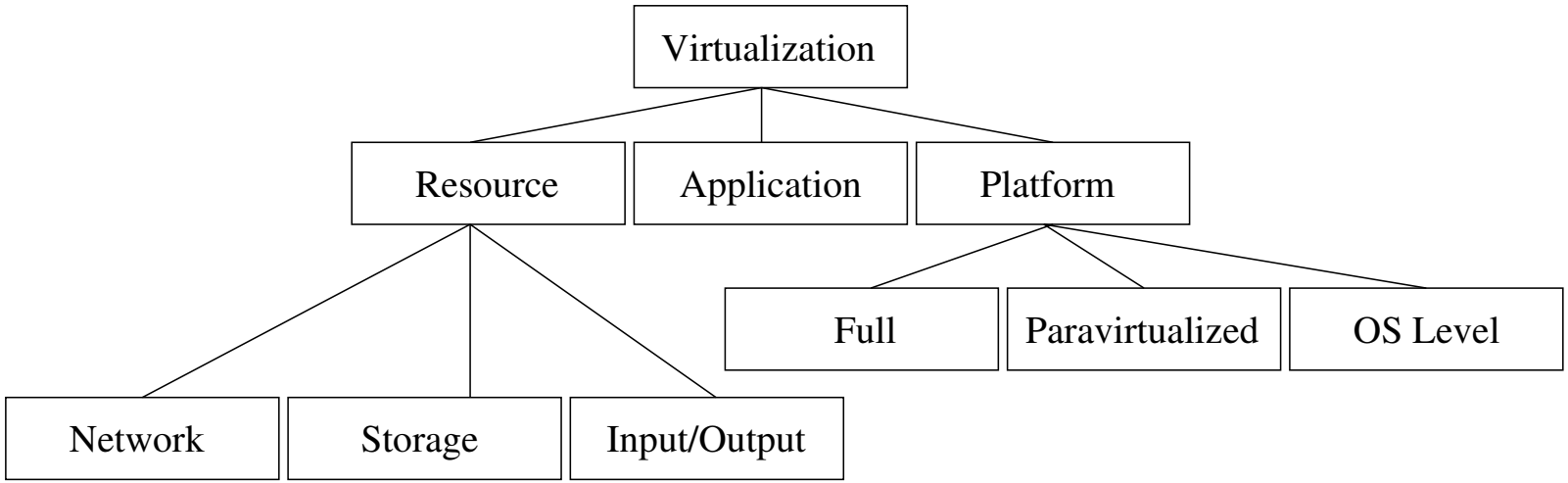
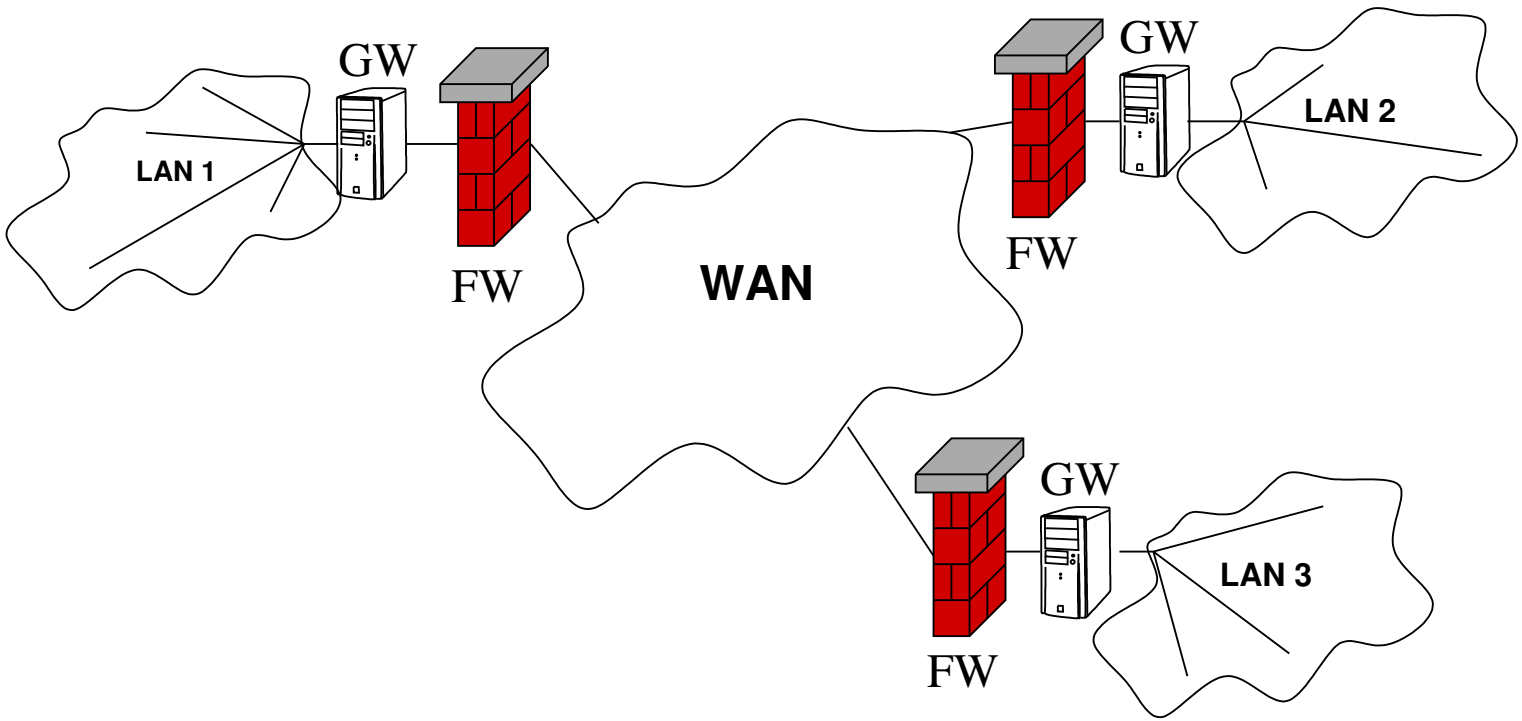


Figure: Types of Virtualization

Network

VPN (Virtual Private Network)

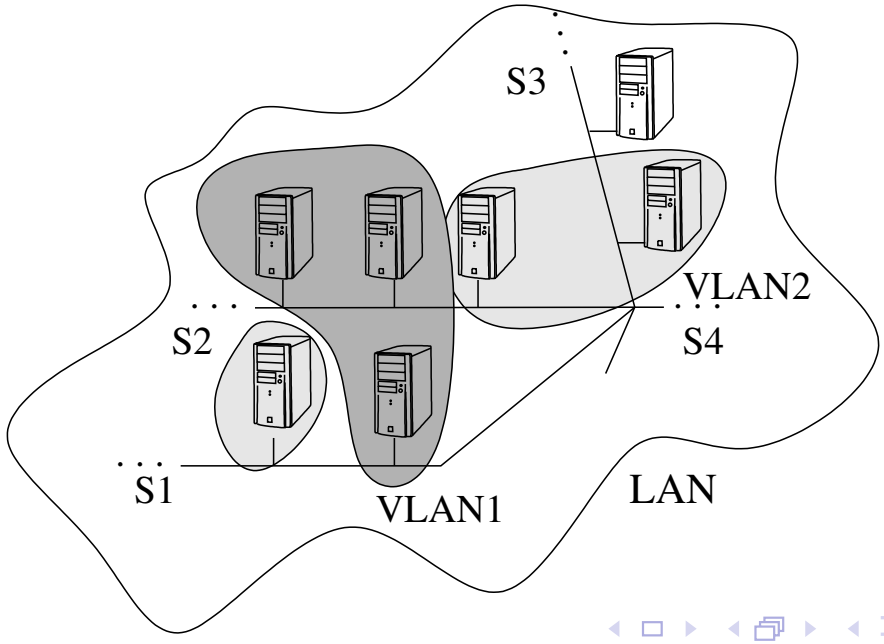
- disjunct network partitions
- gateway service for tunneling
- WAN for interconnect



Network

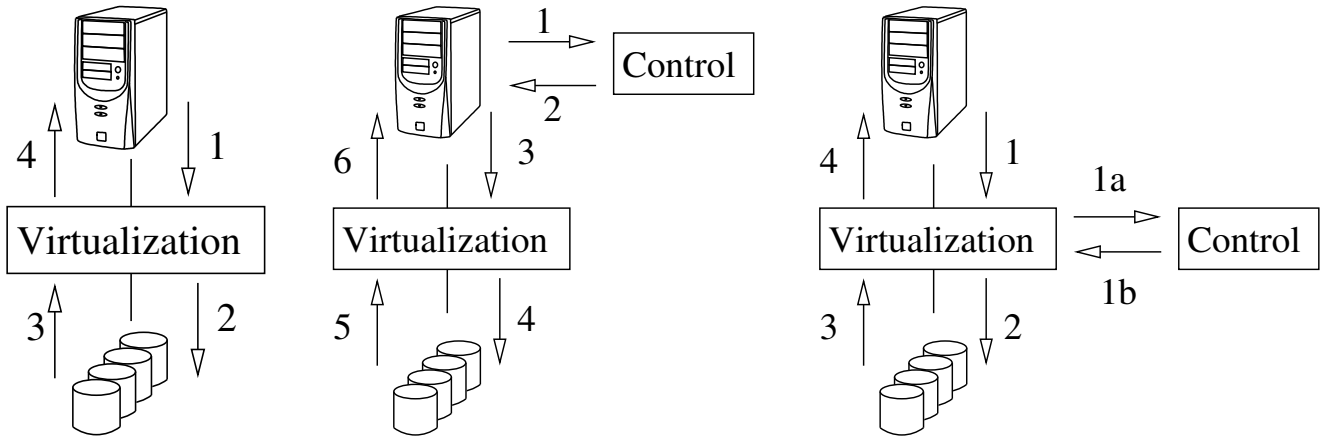
VLAN (Virtual Local Area Network)

- Layer 2 (data link layer) construct, IEEE 802.1Q standard
- Devices on different physical LAN segments
- Communication as if on same physical LAN segment (bcast domain)
- Configuration of VLAN through software
- Traffic shaping/ QoS



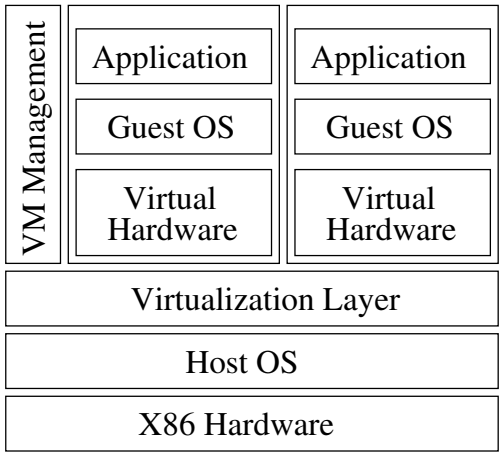
Storage

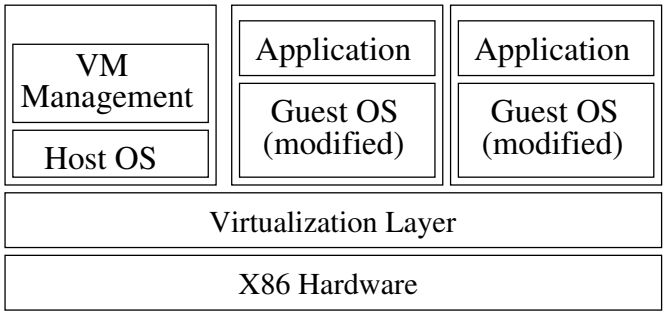
- decreasing storage costs ($\frac{GByte}{\$}$)
- increasing management complexity
- full SAN bandwidth for I/O requests
- separation of data and metadata into different places
- control unit: appliances or SAN switches



Platform

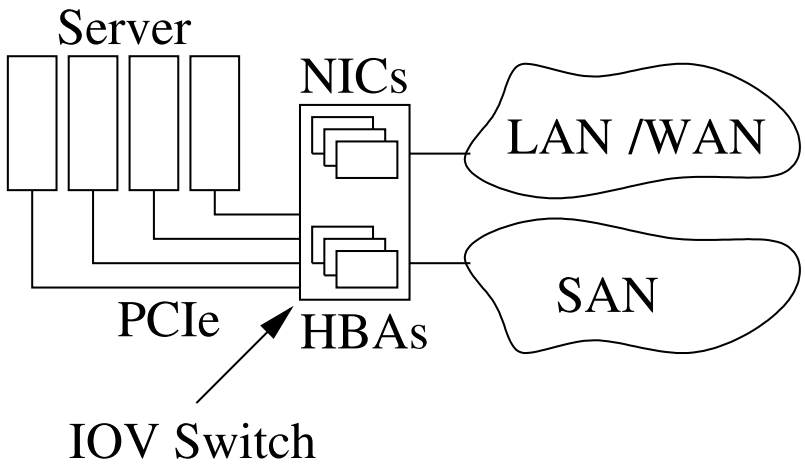
- Full virtualized Environment
  - guest unaware of virtualization, run native OS
  - reduced performance because of hardware emulation
- Para virtualized Environment
  - management module (hypervisor or virtual machine monitor) operates with modified operating system
  - guest OS has much closer control of the underlying hardware (security, influence on other VMs)





Input/ Output (IOV)

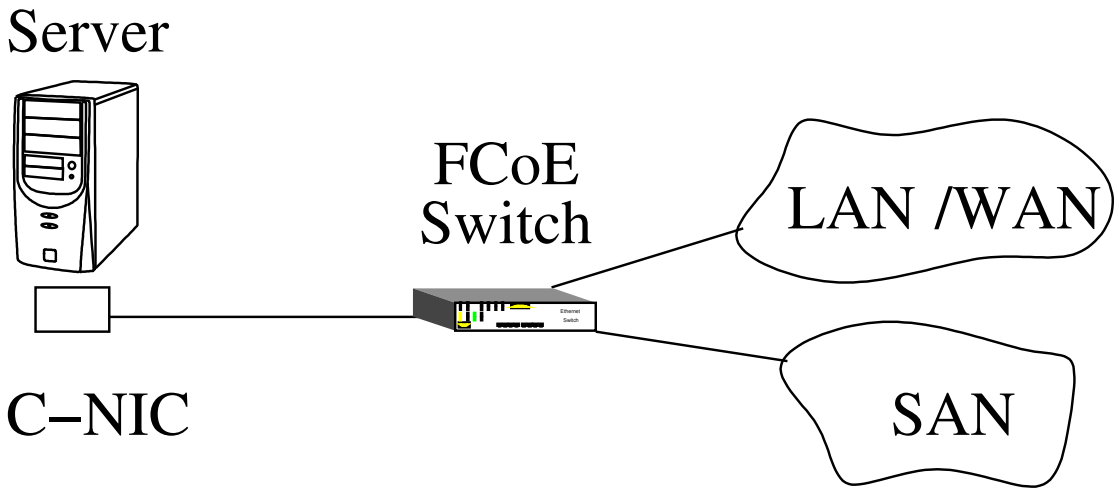
- Typical configuration: 4 to 6 I/O cards per server
  - Ethernet
  - Infiniband
  - Fiber Channel
- add I/O virtualization capability to PCIe
  - disaggreation
  - consolidation
  - virtualization



Input/ Output (IOV)

I/O consolidation

- 100% of server connected to Ethernet, only 20% to FC
- FiberChannel over Ethernet (FCoE), pre-standard
- converged Network Adapter (C-NIC/ CNA)
- benefit for compute centers
  - fewer adapters
  - power savings
  - cable management



Present situation

Platform and storage virtualization in use

- server consolidation
  - improved utilization of existing servers
  - increased number of services per m^2
  - reducing TCO (maintenance,...) in the long-run
- high availability
- run legacy applications
- managing mass-storage backends
- migration of virtual machines is bound to restrictions
  - storage network reconfiguration

Platform Virtualization @ LRMS level

- Job = Virtual Machine
  - share same operations (start, stop, suspend, checkpointing)
- Scenarios:
  - user submits self-prepared virtual machine
  - on-demand creation of virtual machine by LRMS
- VM description as part of job specification
- LRMS schedules jobs and/ or virtual machines

Not so good:

- black-box execution
- performance loss (MPI) compared to non-virtualized solution

IOV @ LRMS level

- vNIC (virtual Network Interface Card)
- vHBA (virtual Host Bus Adapter)
  - FC WWN used to identify to segregate devices for access control
  - Virtualized environments: adapter's WWN represents all DomUs (Xen)
  - DomU migrates to new server, but WWN does not zoning and LUN masking (access control) ineffective
  - vHBA gets own WWN which migrates with VM

Platform Virtualization @ Grid Level

Benefits

- provide uniform environment to Grid users
- rapid on-demand provisioning of Grid nodes
- non-interfering execution of multiple Grid middlewares
- decoupling of knowledge (grid operator/ grid user)
- job exchange, e.g. inter cluster/ Grid resources

Platform Virtualization @ Grid Level

Drawback

- resources run different virtualization software
  - Xen, KVM, VMware, ...
- creation of user environment for each resource

Virtual Appliances

- encapsulate user environment in a virtual appliance

$$\begin{aligned} \text{virtual appliance} &= \text{virtual machine} + \text{operating system} \\ &+ \text{application level software} + \text{MD5} \end{aligned}$$

- Job = ~~virtual machine~~ virtual appliance
- virtual appliance is already technology-dependent :-(
- better approach
  step 1: create

  $$\text{operating system} + \text{application level software} + \text{MD5}$$

  step 2: make it technology-dependent

KIWI

- idea: prepare once, run anywhere
- distribution independent, currently only SuSE supported
- centralized image description based on XML
- output formats: xen, vmware, Amazon EC^2, iso, ...
- adoption to Scientific Linux ongoing in Dortmund

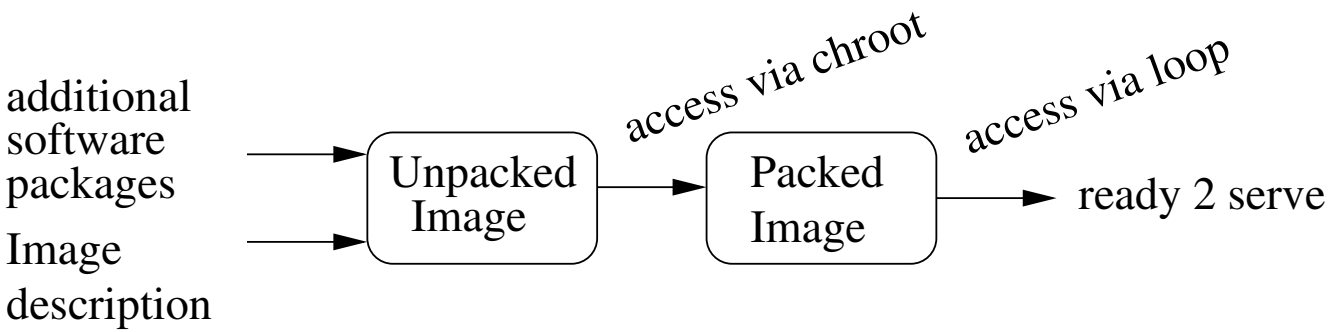


Figure: from: KIWI Cookbook

Conclusion & Outlook

Conclusion

- platform and storage virtualization widely spread
- increased flexibility with I/O virtualization/ FCoE
- Grids profit at LRMS and Grid layer from virtualization
- on-demand provisioning of virtual appliances

Outlook

- Grid resource provides bare metal + virtualization layer
- provision of resources to more than one Grid infrastructure/ VO