Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Virtualisierungskonzepte im Betrieb von gemeinsam genutzten Rechenclustern

Institut für Wissenschaftliches Rechnen
Forschungszentrum Karlsruhe

Institut für Experimentelle Kernphysik
Universität Karlsruhe

Volker Büge, Yves Kemp, Marcel Kunze,
Oliver Oberst, Günter Quast, Christophe Saout

Shared Computing Clusters

A computing cluster shared between different institutes will be built up and maintained by the computing centre of the University of Karlsruhe

Currently 5 participating institutes from different departments

Advantages

- one tender offer leads to more attractive prices
- Central maintenance by few experts
- Homogenous hardware installation
- Intrinsically interception of peak loads by one institute

Problems

- Different requirements of the groups concerning computing infrastructure:
  - Operating system
  - MPI/Network
  - Memory
  - …
- Security aspects
- Sandboxes

Need Virtualisation techniques to really share such a cluster and to enable opportunistic use

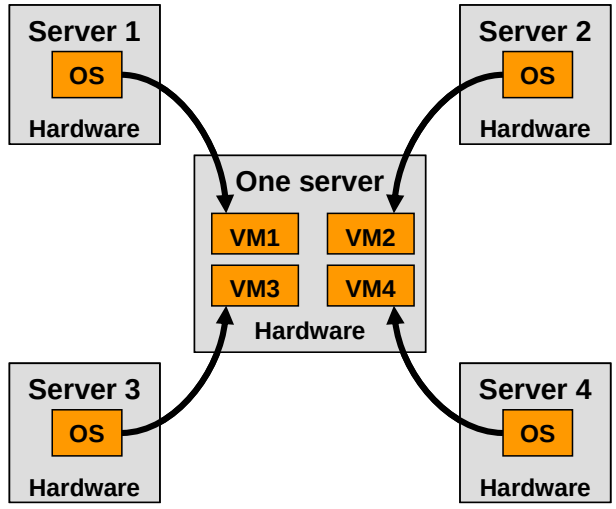
Virtualisation - Concept

Possible definition of virtualisation:

Possibility to share resources of one physical machine between different independent operating systems (OS) in **Virtual Machines (VM)**

Requirements should be:

- Support multiple OS like Linux and Windows on commodity hardware
- Virtual Machines have to be isolated
- Acceptable performance overhead



Virtualisation – Products

There are plenty of existing products for virtualisation:

Win4Lin
Virtual Computing From Desktop to Data Center

Xen™
The Xen™ virtual machine monitor

Linux vServer

vmware®
AN EMC COMPANY

QEMU
open source processor emulator

parallels

OpenVZ

Microsoft
Microsoft Virtual Server

and many more …

Advantages through virtualisation I

Load Balancing

- Only few high performance machines needed to host many different systems
- Deploying a new server means to create a new VM on existing hardware
- Only ordering new hardware if existing is fully loaded
  - → Better distribution of peak usage
  - → Reduction of hardware overhead
  - → cheaper and easier to maintain

Independence of the host OS:

- Free choice of host OS
- Possibility to run different guest OS on each machine
- Virtual machine is decoupled from physical host

→ More flexible administration as requirements concerning the OS of different groups can be satisfied easily

Advantages through virtualisation II

Backup and Cloning:

- Easy backup of whole OS by archiving VM image files
- Cloning of VMs before upgrades enable tests and easy roll back

→ Fewer service interrupts and a more effective administration

Easy and fast setup of new machines:

- Often identical installations of basic OS are needed
- Offer each user group an OS adapted to their needs

→ Deployment of a new machine by duplicating VMs image files

Migration “on the fly”:

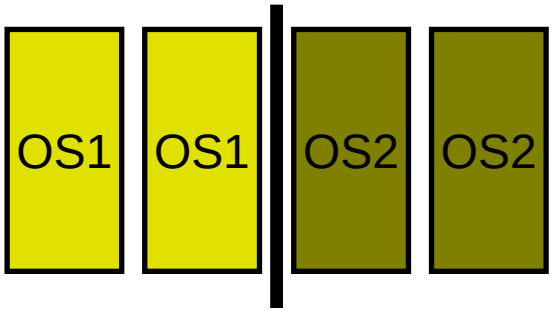
- Possibility of migrating VMs to other host machines
- If the VM file system is stored on a shared storage

  → Migration without interrupt
- High availability solutions

Application - Dynamic Cluster Partitioning

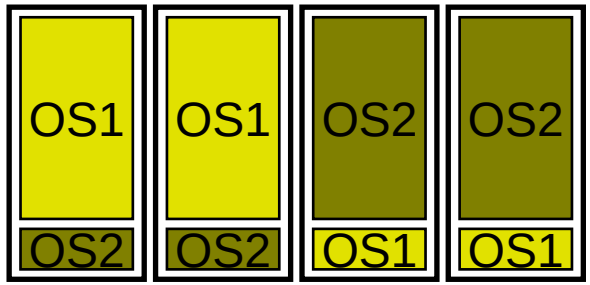
Batch system with 2 different OS required by user groups

Static batch system:



- No resource sharing possible
- Static partition of the cluster
  → Changing partitioning difficult

Virtualised batch system:



- A VM on a free WN can be prepared automatically for an incoming job
- Dynamic partitioning of a cluster possible on short timescales

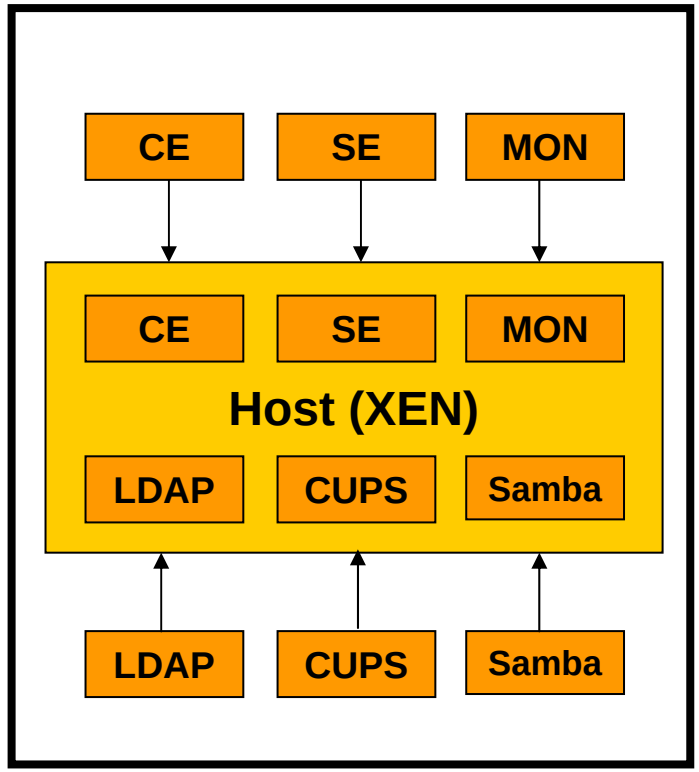
→ Oliver Oberst, T606.2

Applications - Hardware Consolidation I

Typical situation at a university cluster:

- for reasons of stability different services like LDAP, the grid portals, … should run on different machines
- varying load on the different machines
  - → Resources not fully exploited
  - → "recycling" of older machines leads to a heterogeneous hardware structure

→ high administrative effort for installation and maintenance of the system



Virtualisation of these machines leads to few machines to be maintained and to homogenous OS installations

Applications - Hardware Consolidation II

… but what happens if the host machines dies?

- Failure of: disks, motherboard, memory, power supply, …
  - All services which are hosted on this machine will be down until machine is restored or access to VM images possible
- Need concepts of high availability and quality of service for such scenarios where several services are hosted on one physical host

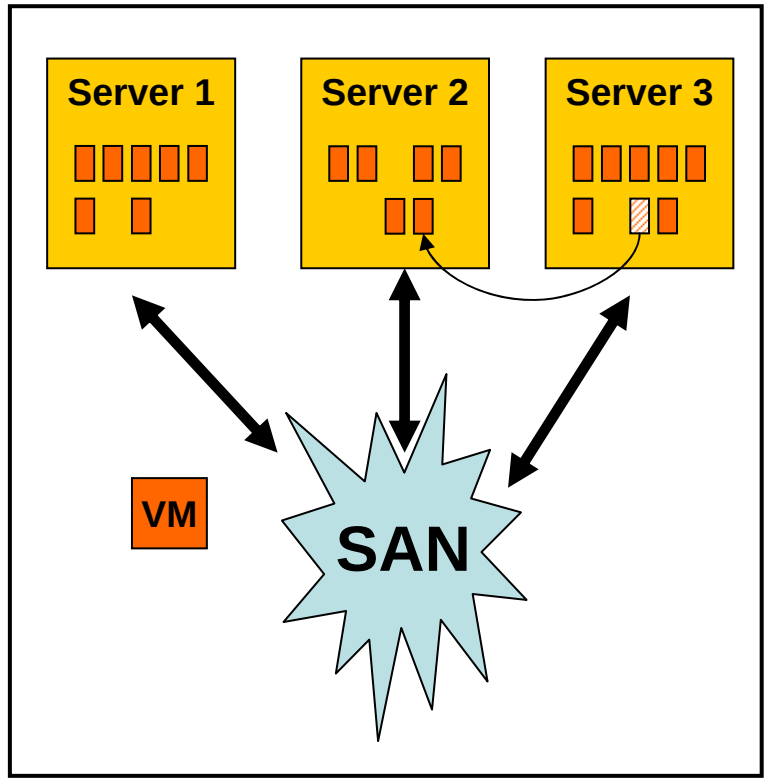
Which techniques can be used to become independent from hardware failures of the host machines?

High available services I

One approach:

- Storage of the VM file system in a high available and redundant SAN
- Use host systems with redundant LAN, SAN and power connections

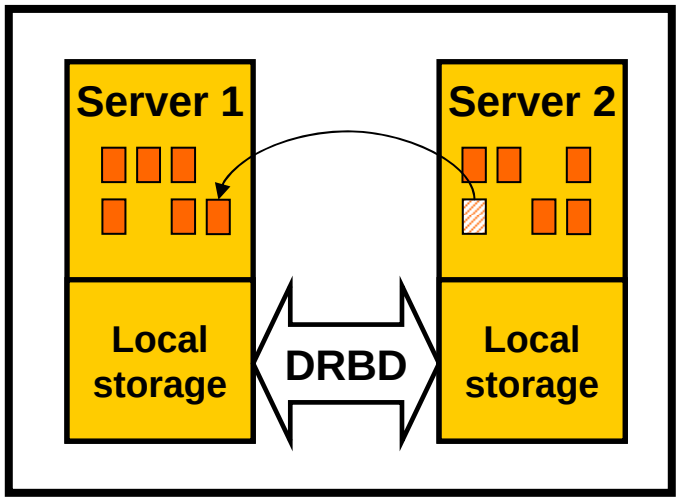
- Migration on the fly in case of hardware problems or maintenance of one server
- If insufficient resources are available on the other server, the service level of less critical services can be reduced for short times.
- Automated tools for load balancing and migration in case of failures exist, e.g for the VMware ESX server.
- All services can be offered without or with only short interruption, perhaps at lower service level



High available services II

Other possibility, e.g for smaller institutes (example EKP):

- Combination of spare machine and SAN overkill if only few critical services are hosted
- Need realisation without too much hardware overhead
- Possibility: Use two performant host machines with same processor architecture and a *Distributed Replicated Block Device* (DRBD) to mirror disk space between both machines



- Local storage containing VM file systems are mirrored on both servers.
- In case of hardware problems on one server, the VM can easily be migrated or restarted on the other

High available services III

Realisation at the IEKP:

Two host machines:

- AMD Athlon(tm) 64 X2 processor
- 6 GB RAM
- 400GB of RAID10 disk space for VMs
- Currently 12 VMs hosted

- Disk space is mirrored on both machines using DRBD
- DRBD via dedicated network interface
- Both machines not fully loaded with VMs to be able to host VMs from other host for short time

Virtualisation allows hardware consolidation combined with high available and redundant services

Conclusion & Outlook

- Building up shared computing cluster together with other institutes has several advantages
- The problems of such shared clusters can be solved using virtualisation techniques
  - Virtualisation of the batch queuing system offers all groups the needed OS but keeps the possibility of opportunistic use
  - It allows the consolidation of many servers on few host machines
    …
- … but we need concepts in case of failures of such a host machine
  - Spare machines and high available SAN for larger computing centres
  - DRBD can be used to build up a high availability infrastructures for a limited number of VMs
  - Concept currently under investigation at the IEKP