High availability clustering of virtual machines – possibilities and pitfalls

Werner Fischer (wfischer@thomas-krenn.com),
Christoph Mitasch (cmitasch@thomas-krenn.com)

Linuxtag 2006 Wiesbaden, Germany
May 6th, 2006

Agenda

1. Problems of virtualization
2. Types of virtualization
3. HA cluster configurations in virtual environments
4. Examples

1) Problems of virtualization

- Setup A: traditional setup (no virtualization)

web-server

db-server

mail-server

file-server

- Setup B: virtualized setup

web-server
db-server
mail-server
file-server

What is the difference between one physical server failing in setup A and one physical server failing in setup B?

In setup B four servers are down :-(

1) Problems of virtualization

"By combining virtualization and HA clustering, it is possible to benefit from increased manageability and savings from server consolidation through virtualization without decreasing uptime of critical services."

James Bottomley in LinuxWorld Magazine (http://linux.syscon.com/read/183014.htm)

2) Types of virtualization

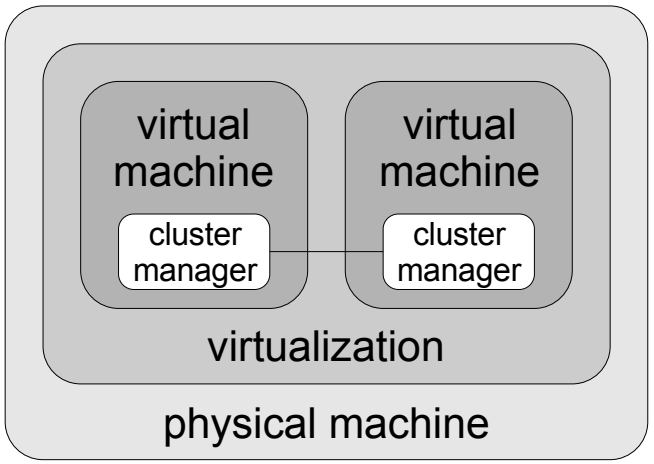
- Hardware-Virtualization (e.g. QEMU, VMware)
  - provides virtual hardware
  - guest OSs run without modifications
- Para-Virtualization (e.g. Xen, Denali, User Mode Linux)
  - today guest OSs (Kernels) need modifications
  - no changes of application binary interface (ABI) needed
- OS-Virtualization (e.g. OpenVZ, Linux-Vserver)
  - only one single kernel runs for both the host and all guests
  - very resource-efficient, guests can be booted in seconds

3) HA cluster configurations

- We will look into these configurations:
  - a) virtual/virtual, single physical machine
  - b) virtual/virtual, multiple physical machines
  - c) physical/virtual
  - d) virtual/physical
  - e) physical/physical with fail-/switchover of virtual machines
- We will also look at:
  - possible pitfalls
  - future

3) HA cluster configurations

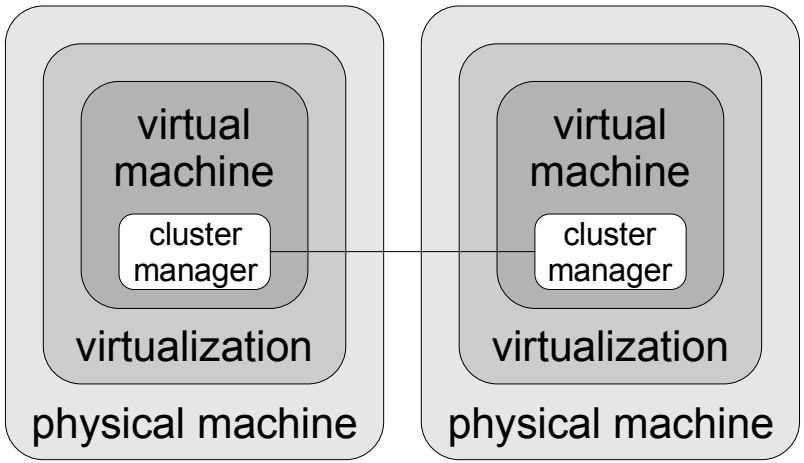
a) virtual/virtual, single physical machine



- does not protect against hardware failures (SPOF)
- protects against kernel panics (hardware- or para-virtualization)
- useful for test environments

3) HA cluster configurations

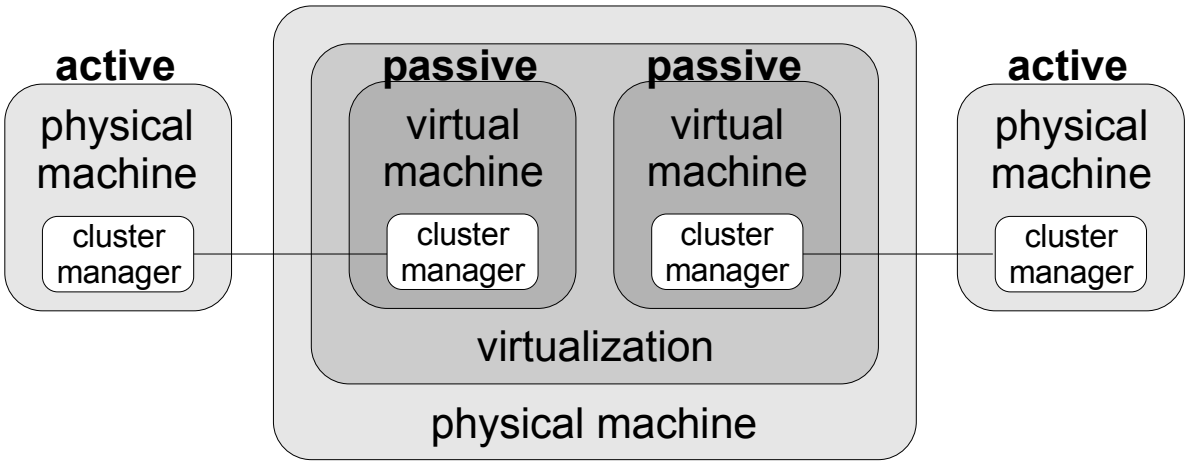
b) virtual/virtual, multiple physical machines



- protects against hardware failures
- easier management of cluster nodes through benefits of virtualization (backup and recovery procedures)
- allows to run multiple instances of an application that normally cannot run multiple times in a single OS instance

3) HA cluster configurations

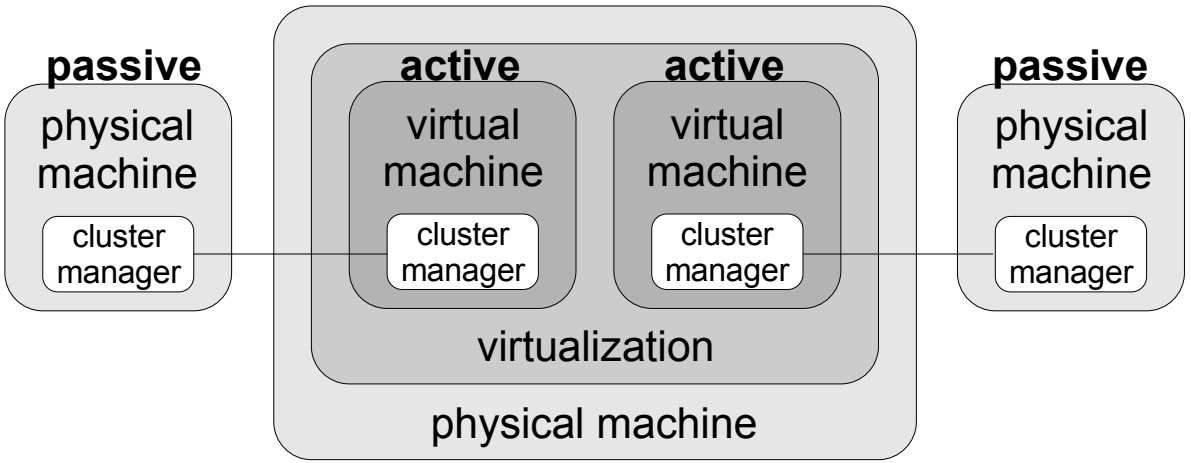
c) physical/virtual



- allows cost savings as only one physical machine can be used as “passive node appliance” for multiple clusters
- useful when the different clusters are running different OSs (hardware/para-virtualization)

3) HA cluster configurations

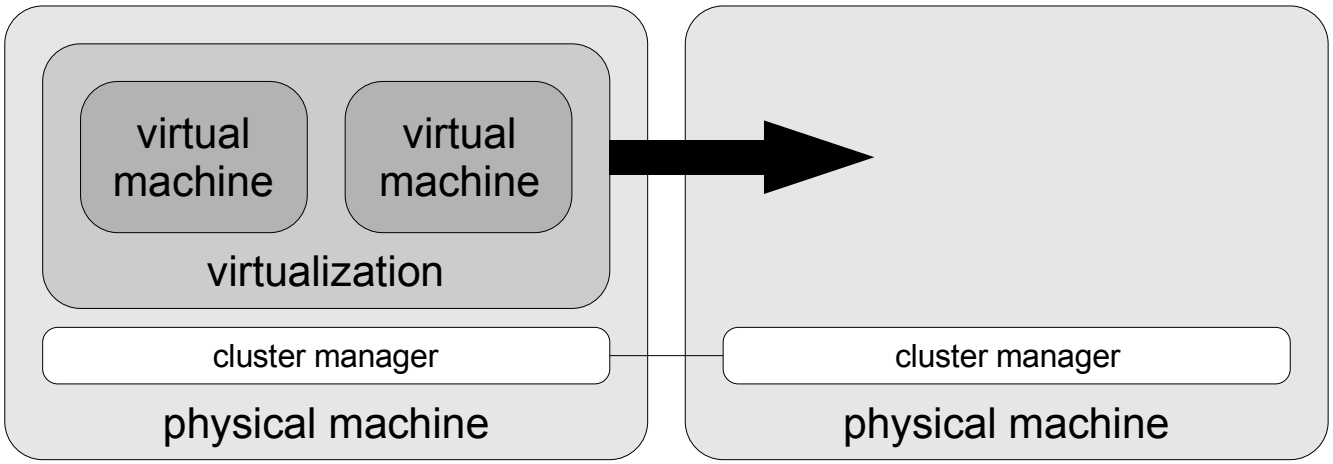
d) virtual/physical



- quite similar to “physical/virtual”, but active and passive roles are interchanged
- “active node appliance” has to be sized large enough
- passive nodes can be less powerful machines

3) HA cluster configurations

e) physical/physical with fail-/switchover of virt. machines



- cluster manager is directly running on the physical machine
- simple cluster configuration – only one HA resource
- no need of cluster-specific changes for HA applications
- very well suited with OS-virtualization (boot time of virtual machines)

3) HA cluster configurations

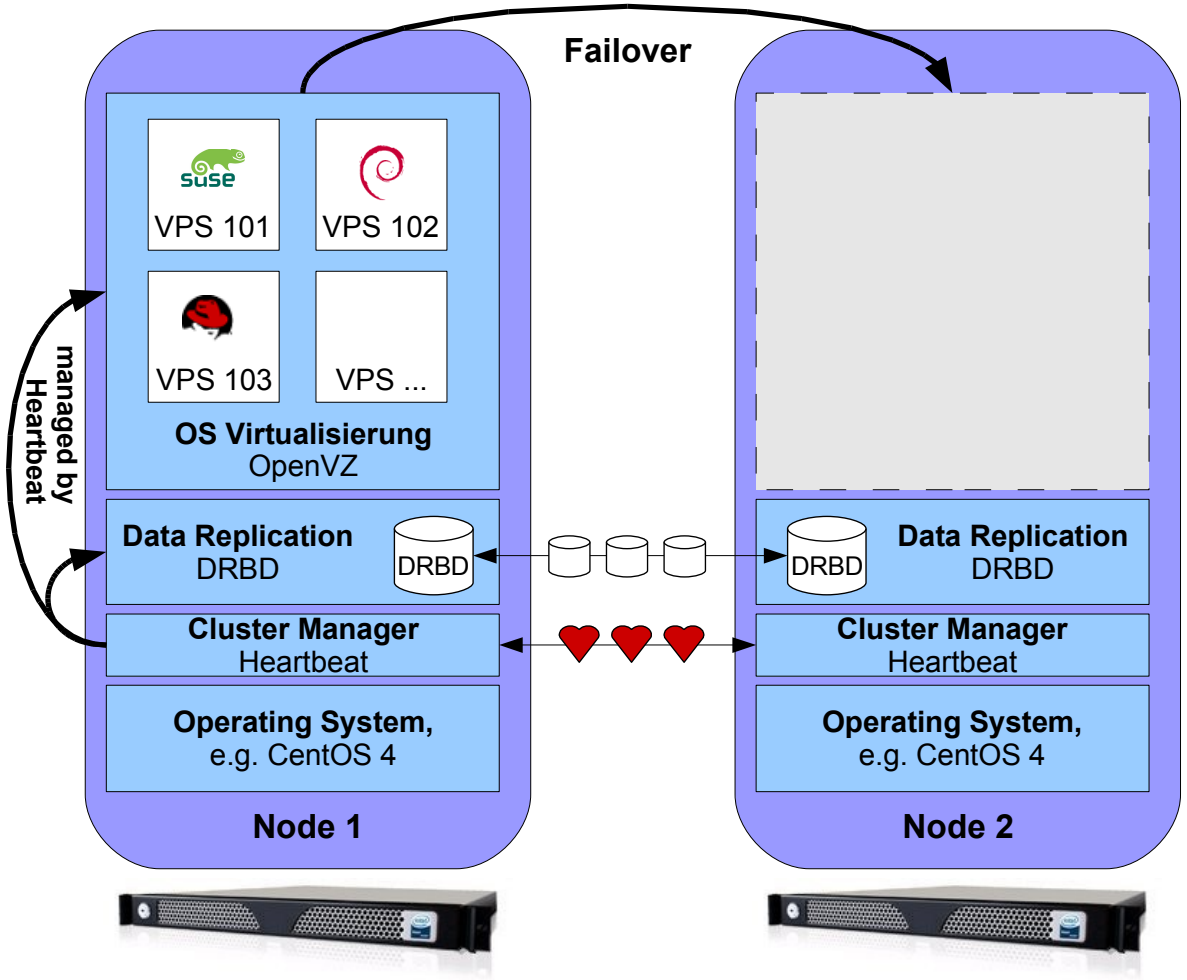
- possible pitfalls
  - HA clustering solutions have not been built with virtualization in mind
  - especially shared resource protection mechanisms can cause troubles, e.g.
    - resource fencing (SCSI-2/SCSI-3 reservations)
    - node fencing (STONITH – shoot the other node in the head)
  - remember the key HA principle: simplicity

3) HA cluster configurations

- future
  - combining virtualization and HA clusters is just in its beginning stages at the moment
  - real-time migration is currently getting very popular for virtualization technologies (allows to migrate a virtual machine to another server while preserving memory, processes and network connections)
    - this feature could be used for switchovers in an HA cluster
    - perhaps some of this ideas could be also used for failover situations?

4) Examples

- Cluster with OpenVZ, Heartbeat, and DRBD



4) Examples

- Cluster with Linux-Vserver, Heartbeat, and DRBD: http://linux-vserver.org/Vserver+DRBD
- RedHat Cluster Suite inside a Xen Guest: http://people.redhat.com/pcaulfie/docs/xencluster.html
- Heartbeat 2.X resource agent for Xen guests: http://lists.linux-ha.org/pipermail/linux-ha/2006-March/018798.html

What I've been talking about...

1. Problems of virtualization
2. Types of virtualization
3. HA cluster configurations in virtual environments
4. Examples

Thanks for your interest in this topic

And also thanks Thomas-Krenn.AG for letting Christoph and me work on this talk and the paper at work :-)