# High-Availability Using Open Source Software

Luka Perkov, Nikola Pavković, Juraj Petrović

Iskon Internet d.d.
Garićgradska 18, 10000 Zagreb, Croatia

Ruđer Bošković Institute
Bijenička cesta 54, 10000 Zagreb, Croatia

Faculty of Electrical Engineering and Computing, University of Zagreb
Unska 3, 10000 Zagreb, Croatia

May 26, 2011

# Outline

# Overview

Computer clusters can be categorized into three groups:

- High Availability clusters (HA)
- Load Balancing clusters (LB)
- High Performance Computer clusters (HPC)

**HA cluster goal**

If the event of complete node failure occurs, the service provided by a HA cluster should continue to operate normally and recover to the initial state successfully upon node recovery.

# HA architecture overview

When designing a HA system one should use a combination of loosely coupled components to achieve the desired goal. These components are called resources and could be perceived as HA cluster building blocks.

Three main resources, where HA implementation would be advisable, are:

- storage
- application
- operating system

# Storage resources overview

Here described technologies are presented as *virtual* storage block devices to the host OS. This setup makes it possible to choose from variety of file systems that can be used on any storage device.

Mostly used are:

- DRBD
- Open-iSCSI and iSCSI Enterprise Target

# DRBD overview

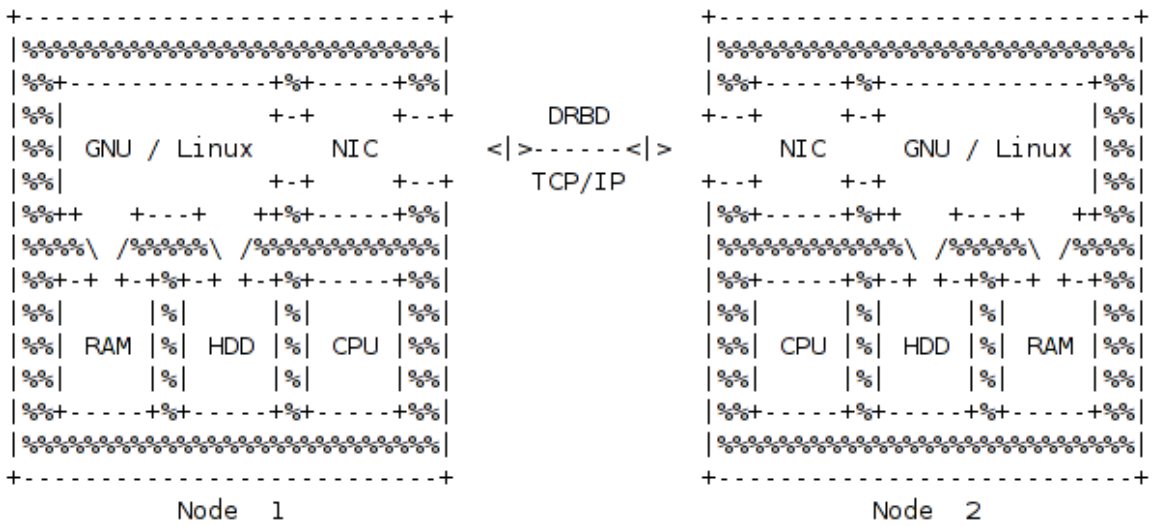
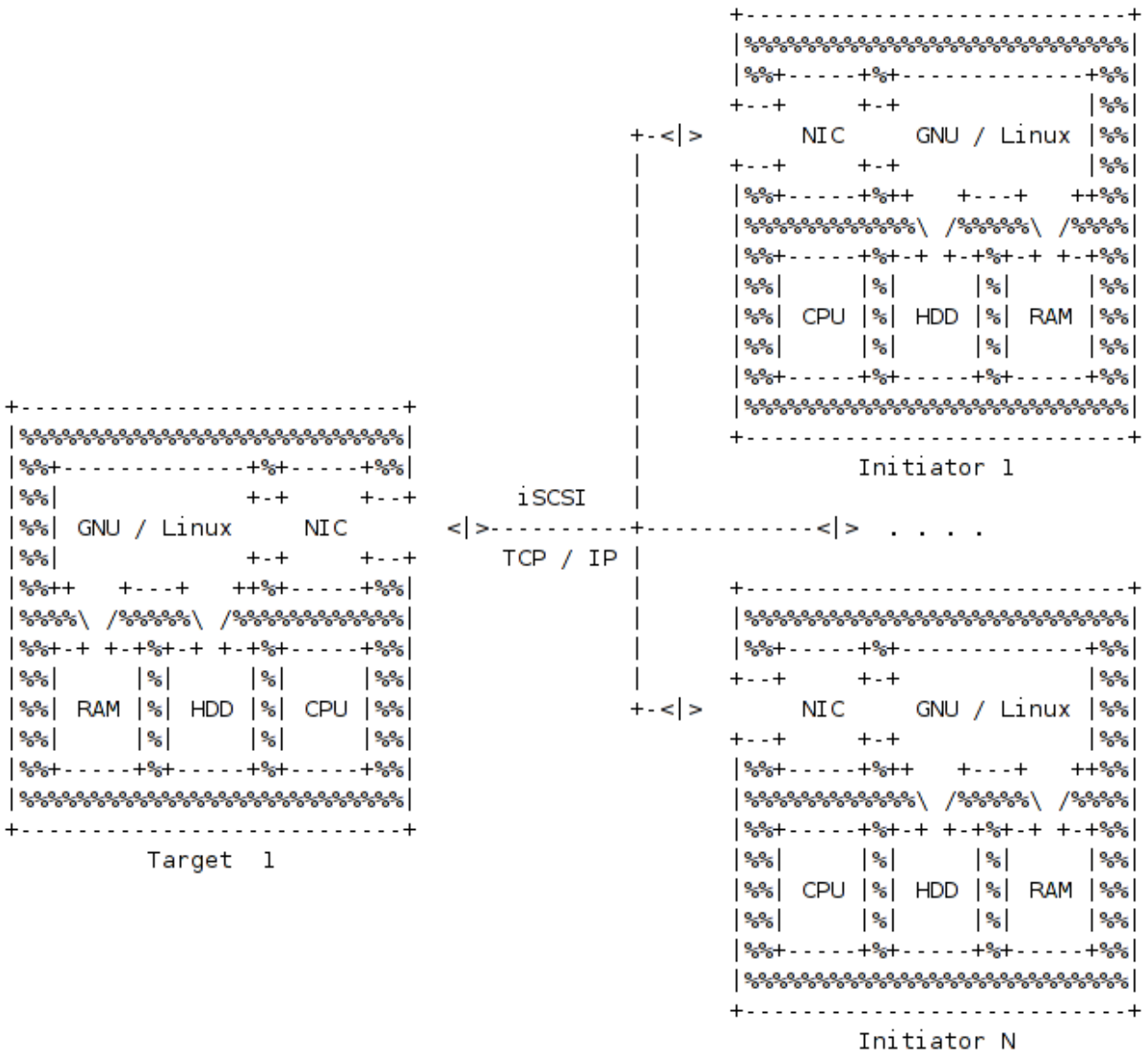



Figure: DRBD architecture

# iSCSI overview



# OS virtualization overview

OS virtualization must be used in order to provide full OS HA.

One of the features that all of described solutions provide is *live* migration, which is convenient when migrating virtual machines (VM) between physical nodes.

Some of the available OS virtualization solutions:

- Xen
- KVM
- OpenVZ

# Xen overview

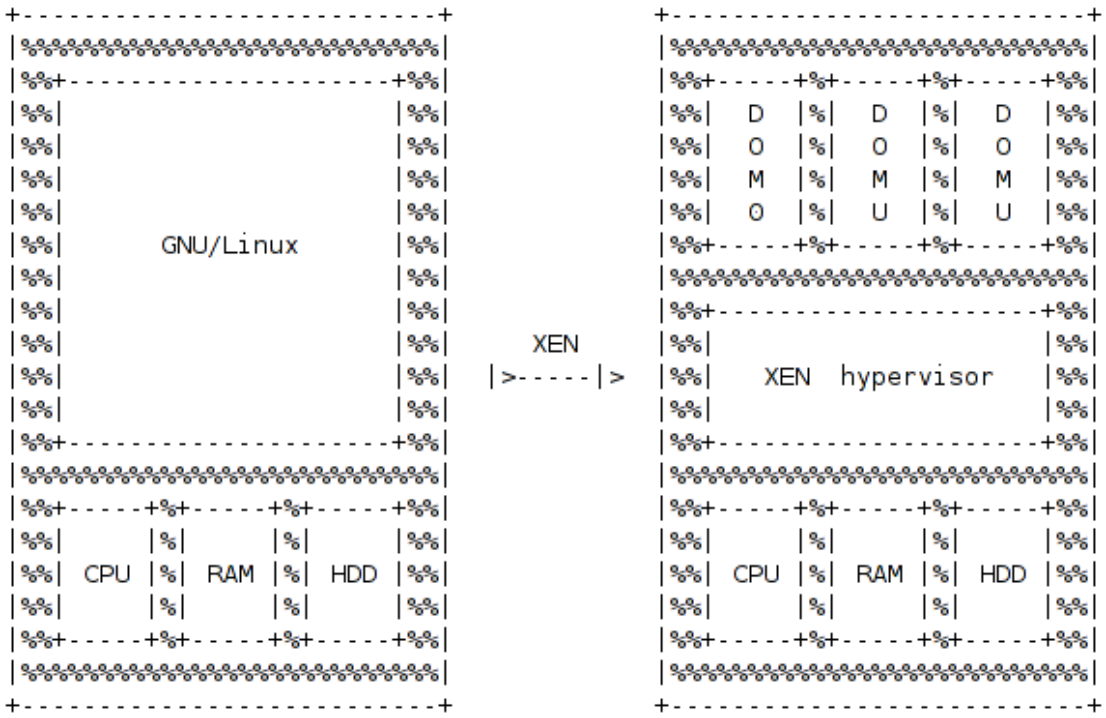



Figure: Xen architecture

# KVM overview

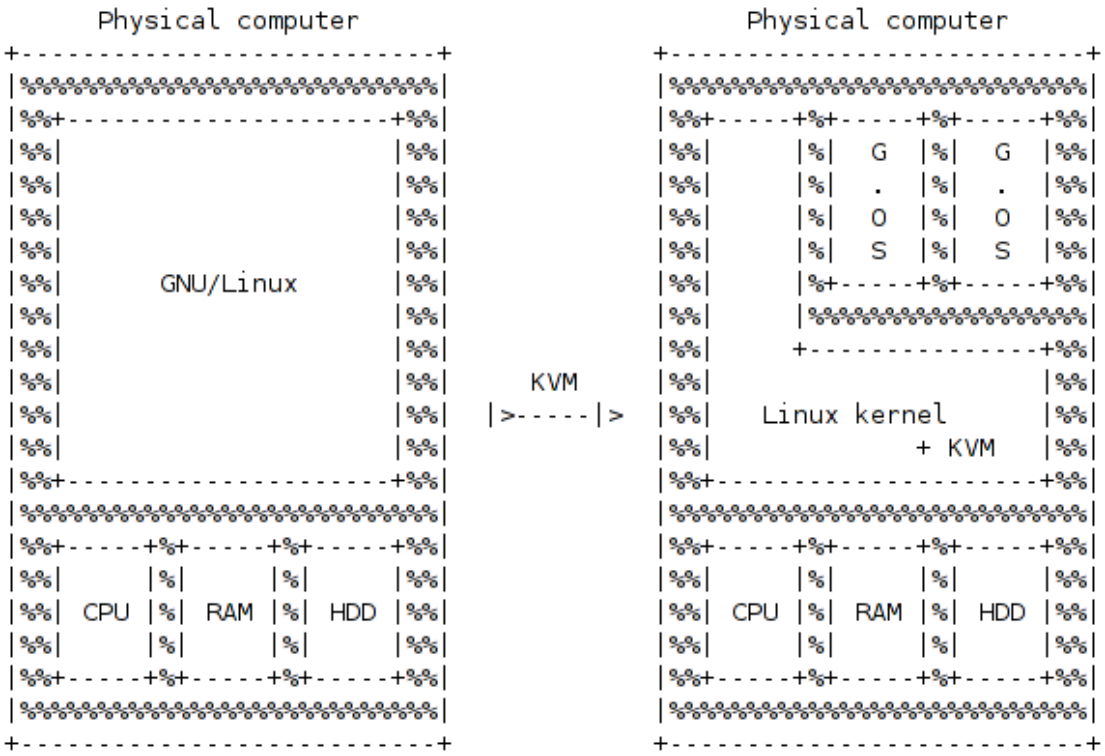



Figure: KVM architecture

# OpenVZ overview

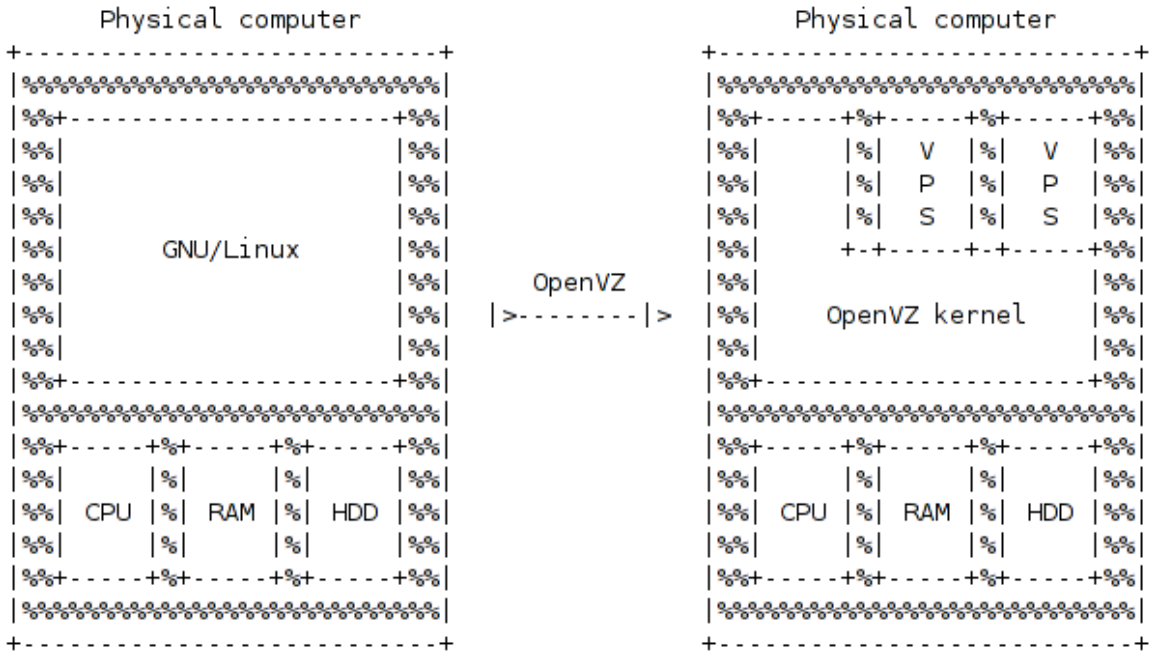



Figure: OpenVZ architecture

# Cluster management overview

In order to fully automate recovery by detecting hardware or software failures it is essential to use appropriate tools which provide advanced cluster node and resource management operations.

Only when adequate HA cluster resource management (CRM) tools have been deployed, cluster management is both adjustable and extensible.

# Pacemaker overview

Pacemaker CRM solution makes no assumptions about desired environment, which allows to provide support for practically any redundancy configuration including Active/Active, Active/Passive, N+1, N+M, N-to-1 and N-to-N.

At the highest level, pacemaker cluster architecture consists of three pieces:

- core cluster infrastructure providing messaging and membership functionality (eg. Corosync or Heartbeat),
- non-cluster aware components and
- a brain, processing the events from the cluster and configuration changes.

# Conclusion

Using a variety of Open Source tools it is possible to develop and maintain fully functional HA cluster that meets the requirements of business.

If carefully deployed, Open Source HA cluster will have equal, if not better performance of its commercial alternatives on the same hardware platform.

In the long term, time invested in research, development and testing custom cluster deployments has more benefits than purchasing expensive of the shelf cluster solutions.