

**VirtualLinux** is a Linux meta-distribution that allows the creation, deployment and administration of virtualized clusters with no single point of failure. VirtualLinux architecture supports diskless configurations and provides an efficient, iSCSI based abstraction of the SAN. Clusters running VirtualLinux exhibit no master node, thus boosting resilience and flexibility.

### DISKLESS CLUSTER + EXTERNAL SAN

- Using EVMS and iSCSI, the architecture provides a flexible, high-level description of the underlying hardware that frees the administrator from the traditional, rigid allocation of resources.
- Performance and scalability are achieved by means of the direct access of cluster blades to the external SAN via a switched Infiniband network (10-20Gb/s).
- Storage reliability is implemented within the external SAN, which can exploit a redundant array of disks (disks that can be mounted on the blades are usually quite slow and fragile).

### MASTERLESS CLUSTER CONFIGURATION

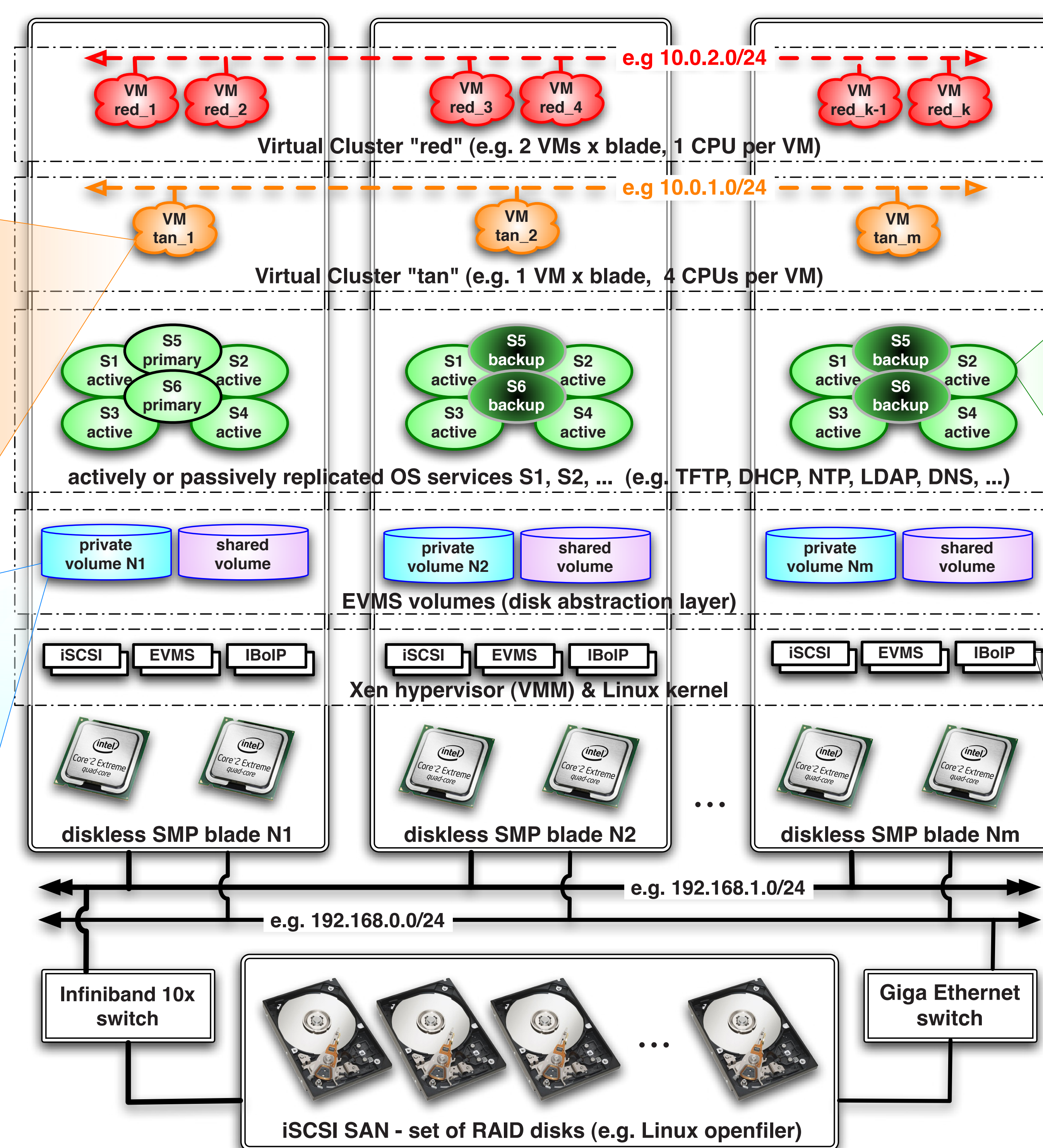
- VirtualLinux has no master, all nodes have a symmetric configuration.
- Critical OS services are categorized and made redundant by either active or passive replication in such a way they are, at each point in time, cooperatively implemented by the running nodes.
- Any blade of the cluster can be hot-swapped with no impact on cluster operation.

### CLUSTER VIRTUALIZATION

VirtualLinux separates three environments, targeted to different classes of administrators:

- The physical cluster, including the physical devices, which are insulated and made transparent (hardware technician).
- The privileged cluster, i.e. the environment that provides services and interfaces to the virtual clusters (skilled OS administrator).
- The virtual clusters that are sets of virtual machines. They can run any OS and configuration in such a way Grid and Beowulf style virtual clusters can co-exist simultaneously (standard OS administrator).

### VIRTUALINUX ARCHITECTURE



#### VIRTUAL CLUSTERS (VC)

- Each virtual node (VM) of a VC is a virtual machine that can be configured at creation time. It exploits a cluster-wide shared storage.
- Each VC exploits a private network and can access the cluster external gateway.
- VMs of a VC can be flexibly mapped onto the cluster nodes.
- VCs can be dynamically created, destroyed, suspended on disk.

#### LINUX OS SERVICES

All standard Linux services are made fault-tolerant via either active or passive replication:

- Active: Services are started in all nodes; a suitable configuration enforces load balance on client requests. E.g. NTP, DNS, TFTP, DHCP.
- Passive (primary-backup): Linux HA with heartbeat is used as fault detector. E.g. LDAP, IP gateway.

#### DISKS ABSTRACTION LAYER

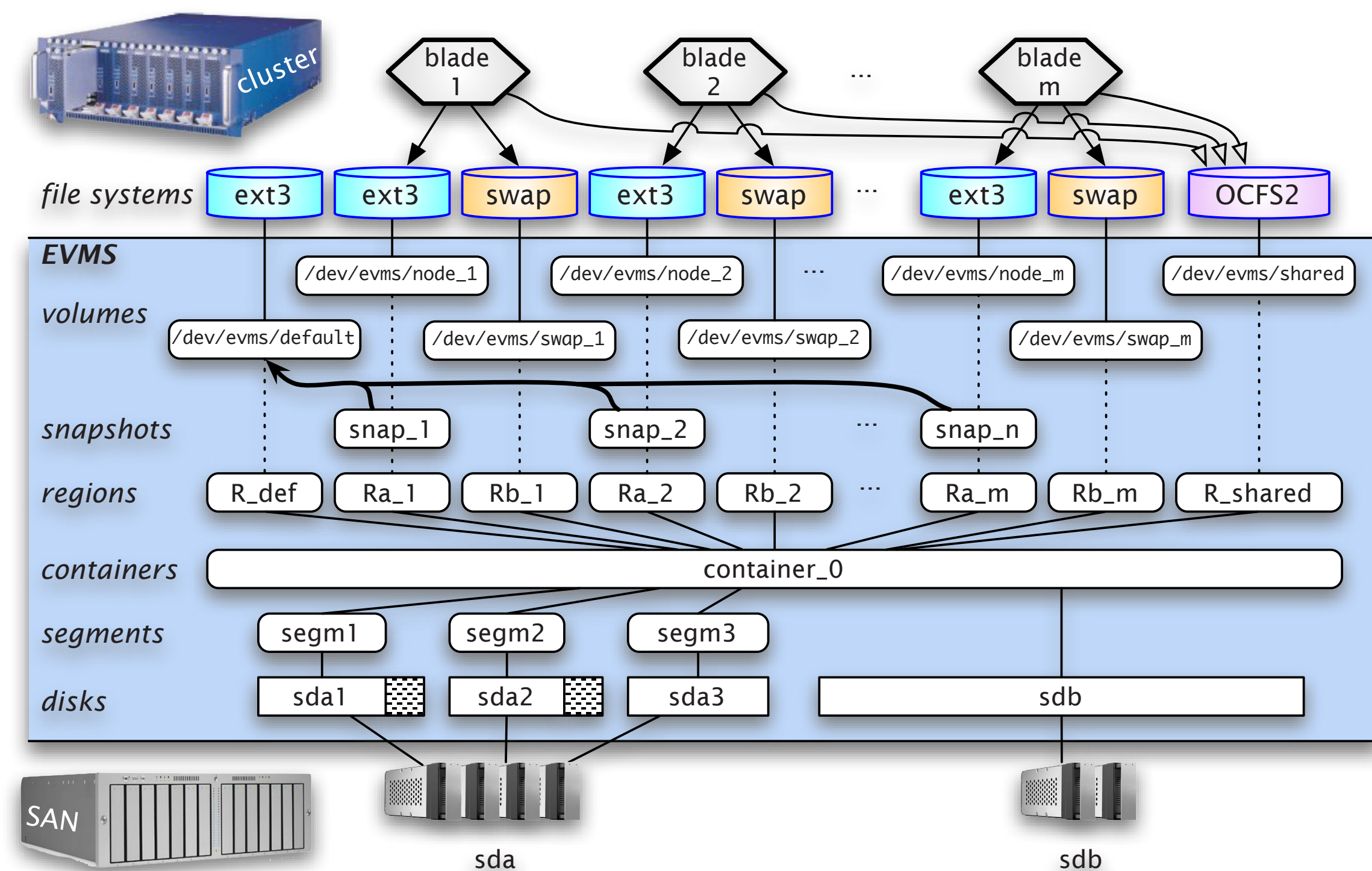
A set of private and shared EVMS volumes are mounted via iSCSI in each node of the cluster:

- A private disk (/root) and a OCFS2/GFS cluster-wide shared SAN are mounted in each node.
- EVMS snapshot technique is used for a time and space efficient creation of the private remote disk.
- A novel plug-in of EVMS has been designed to implement this feature.

#### KERNEL BASIC FEATURES

- All standard Linux modules.
- Xen hypervisor, supporting Linux paravirtualization, and Microsoft Windows via QEMU binary translation (experimental).
- Network connectivity, including Infiniband userspace verbs and IP over Infiniband.
- iSCSI remote storage access.
- OCFS2 and GFS shared file systems.

### DISK ABSTRACTION LAYER ARCHITECTURE



### EUROTECH CLUSTERS EQUIPPED WITH VIRTUALINUX



### VIRTUALINUX FEATURES AND TOOLS

VirtualLinux provides:

- A bootable DVD.
- Distribution independence: Virtual Clusters can run any distribution (currently included virtual clusters: Ubuntu Edgy 6.10 and CentOS 4.4 for x86\_64).
- An install facility to setup and configure the included virtual machine images in a diskless, masterless fashion.
- A recovery facility able to reset a misconfigured (physical) node to factory status.
- A toolkit to manage Virtual Clusters, which can be dynamically created, destroyed and moved.
- User and developer documentation.

### SUMMARY

VirtualLinux allows the coexistence of many Linux distributions and configurations in the same cluster; moreover, virtual clusters can be saved, restored and moved. Upgrading and testing of the cluster is greatly simplified and requires less time. In addition, fault-tolerance is obtained by a combination of architectural, software and hardware strategies that implement a design with no single point of failure.

### DESIGNERS, DEVELOPERS, AND CREDITS

University of Pisa: Marco Aldinucci, Marco Danelutto, Francesco Polzella, Gianmarco Spinatelli, Massimo Torquati, Marco Vanneschi – Eurotech HPC: Manuel Cacitti, Alessandro Gervaso, Pierfrancesco Zuccato. VirtualLinux has been developed at the HPC lab. of Computer Science Dept. – University of Pisa and Eurotech HPC, a division of Eurotech Group. VirtualLinux project has been supported by the initiatives of the LITBIO Consortium, founded within FIRB 2003 grant by MIUR, Italy.

### REFERENCES

M. Aldinucci, M. Torquati, M. Vanneschi, A. Gervaso, P. Zuccato, M. Cacitti. *VirtualLinux Design Principles*. Technical Report TR-07-13, Computer Science Department, University of Pisa, Italy, June 2007. <http://compass2.di.unipi.it/TR/Files/TR-07-13.pdf.gz>  
M. Aldinucci, M. Danelutto, M. Torquati, F. Polzella, G. Spinatelli, M. Vanneschi, A. Gervaso, M. Cacitti, P. Zuccato. *VirtualLinux: virtualized high-density clusters with no single point of failure*. Intl. PARCO 2007: Parallel Computing, September 2007, to appear.  
The Eurotech Group: <http://www.eurotech.com> – HPC lab. of Uni. Pisa Computer Science Dept.: <http://www.di.unipi.it/groups/architettura/> – LITBIO: <http://www.litbio.org>