PAUL SCHERRER INSTITUT

From Real to Virtual - How to Provide a High-Availability Computer Server Infrastructure

R. Kapeller, R. Krempaska, H. Billich, P. Huesser, C. Higgs, H.Lutz
Paul Scherrer Institute, 5232 Villigen PSI, Switzerland

During the commissioning phase of the Swiss Light Source (SLS) at the Paul Scherrer Institute (PSI) we decided in 2000 for a strategy to separate individual services for the control system. The reason was to prevent interruptions due to network congestion, mis-directed control, and other causes between different service contexts. This concept proved to be reliable over the years. Today, each accelerator facility and beamline of PSI resides on a separated subnet and uses its dedicated set of service computers. As the number of beamlines and accelerators grew, the variety of services and their quantity rapidly increased. Fortunately, about the time when the SLS announced its first beam, VMware introduced its VMware Virtual Platform for Intel IA_32 architecture. This was a great opportunity for us to start with the virtualization of the controls services. Currently, we have about 200 such systems. In this presentation we discuss the way how we achieved the high-level-virtualization controls infrastructure, as well as how we will proceed in the future.

Types of Services/Servers

Service	Total
SSH Gateway	23
Channel Access GW	40
Softioc	37
Port Server Host	23
IOC Boot PC	32
WWW, Elog, Test, ..	35
Total	190

Why to separate Services?

Easier to manage
- Service adapted installation and configuration

Proportioned security
- Make the system as secure as the service requires

No negative service interaction
- An "out of control" service does not harm others

No resource conflicts
- (Almost) exclusive use of RAM, Disk, Network, etc

Reboot without fear
- Untrained personal may restart the system in an emergency case
- A system reboot always applies to one service

Virtualization for the crowd

The start of affordable computer virtualization on IA-32 and IA-64 systems was a great opportunity to introduce virtualization at the Swiss Light Source.

Terms

V12N	**Virtualization** (use of numeronym to be concise)
VM	**Virtual Machine** (a software based, fictive computer)
vFiler	**Virtual File Server** (a software based, fictive storage server)
RHEL	**Red Hat Enterprise Linux** (a Unix-based operating system)
SAN	**Storage Area Network** (dedicated network, that provides access to block level data storage)
FC	**Fibre Channel** (high speed network to connect data storage)

Cluster specs

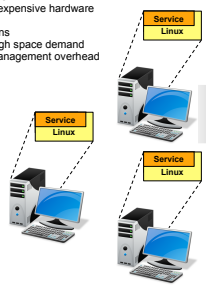
Clusters	2
Nodes	6
Cores	48
Virtual Machines	221
Total GB RAM	512
Total GHz CPU	120
LUNs	9 à 600GB

Service Separation

2002 - Use of one Linux PC per service eg. IOC boot server

Pros
-Inexpensive hardware

Cons
-High space demand
-Management overhead

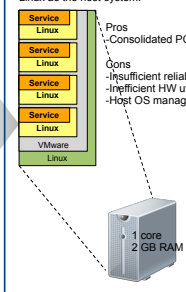


Vmware Server (1)

2004 - Initial virtualization step, with Virtual machines (PC) inside "VMware Server" and Linux as the host system.

Pros
-Consolidated PCs

Cons
-Insufficient reliability
-Inefficient HW utilization
-Host OS management

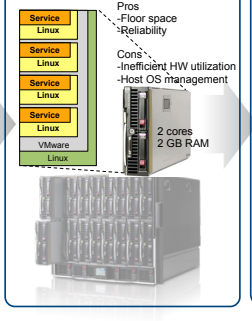


Vmware Server (2)

2006 - Same as previous, but improved hardware reliability and optimized floor space.

Pros
-Floor space
-Reliability

Cons
-Inefficient HW utilization
-Host OS management

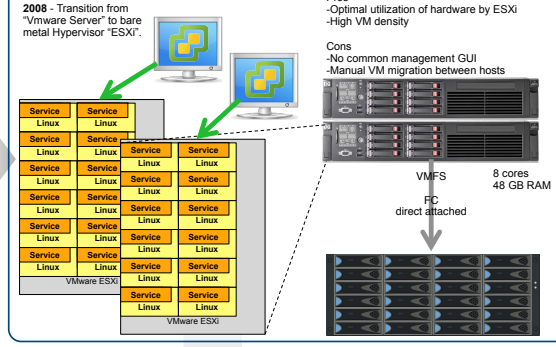


ESXi

2008 - Transition from "Vmware Server" to bare metal Hypervisor "ESXi".

Pros
-Optimal utilization of hardware by ESXi
-High VM density

Cons
-No common management GUI
-Manual VM migration between hosts

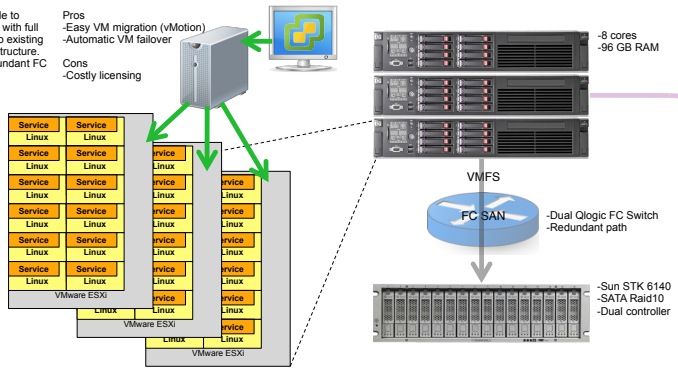


ESXi with vCenter

2011 - Upgrade to licensed ESXi with full integration into existing vCenter infrastructure. Attaching redundant FC SAN.

Pros
-Easy VM migration (vMotion)
-Automatic VM failover

Cons
-Costly licensing

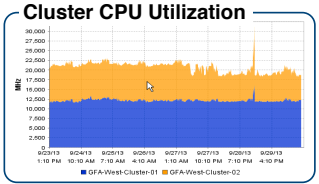


VM Backup

Fast reproduction (installation & configuration) of VM in favor of a costly backup solution.

Using RHEL kickstart, in combination with a central configuration service (puppet) and a defined host class schema, any production VM Server must be re-installable within hours. Therefore we have not scheduled regular backups for our VMs. Occasionally we use 'ghettoVCB.sh' to write a copy of each VM to a dedicated NFS Server. The script takes snapshots of a live running VM, backs up the master VMDK and then upon completion, deletes the snapshot until the next backup.
A backup solution at no cost, that requires SSH access to the ESXi servers and some basic Linux skills.

Cluster CPU Utilization

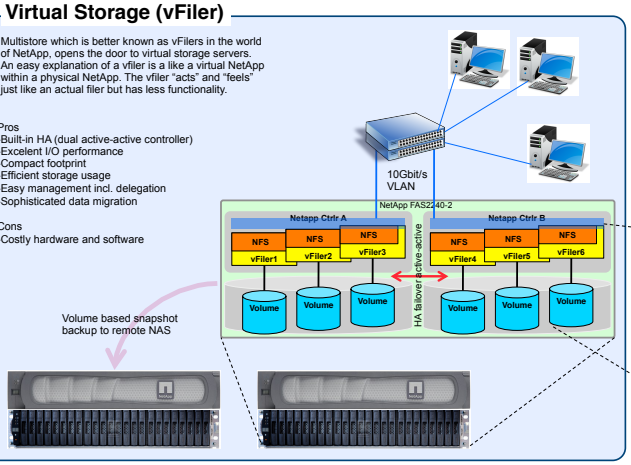


Virtual Storage (vFiler)

Multistore which is better known as vFilers in the world of NetApp, opens the door to virtual storage servers. An easy explanation of a vfiler is a like a virtual NetApp within a physical NetApp. The vfiler "acts" and "feels" just like an actual filer but has less functionality.

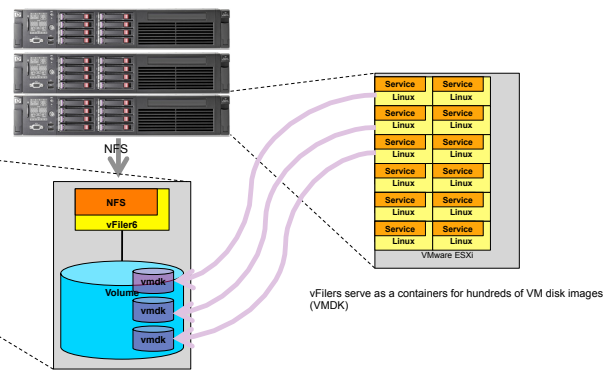
Pros
-Built-in HA (dual active-active controller)
-Excelent I/O performance
-Compact footprint
-Efficient storage usage
-Easy management incl. delegation
-Sophisticated data migration

Cons
-Costly hardware and software



VM disks on vFiler storage

2014 – The current SAN storage infrastructure for the VM disk images will be migrated to NFS virtual storage (vFiler).



vFilers serve as a containers for hundreds of VM disk images (VMDK)