APPLICATION OF VIRTUALIZATION TO CERN ACCESS AND SAFETY SYSTEMS

T. Hakulinen, J.B. Lopez Costa, P. Ninin, H. Nissen, R. Nunes, CERN, Geneva, Switzerland

Abstract

Access and safety systems are by nature heterogeneous: different kinds of hardware and software, commercial and home-grown, are integrated to form a working system. This implies many different application services, for which separate physical servers are allocated to keep the various subsystems isolated. Each such application server requires special expertise to install and manage. Furthermore, physical hardware is relatively expensive and presents a single point of failure to any of the subsystems, unless designed to include often complex redundancy protocols. We present the Virtual Safety System Infrastructure project (VSSI), whose aim is to utilize modern virtualization techniques to abstract application servers from the actual hardware. The virtual servers run on robust and redundant standard hardware, where snapshotting and backing up of virtual machines can be carried out to maximize availability. Uniform maintenance procedures are applicable to all virtual machines on the hypervisor level, which helps to standardize maintenance tasks. This approach has been applied to the servers of CERN PS and LHC access systems as well as to CERN Safety Alarm Monitoring System (CSAM).

INTRODUCTION

Virtualization

Computer virtualization can be defined as emulating existing computer hardware in software in such a way that the operating system and applications running inside that environment are not aware, at least in principle, that they are not being run directly on the hardware platform in question. The machine running the virtualization software is called the *host*, the machine being virtualized is called a *client* or *guest*, and the virtualization software is called a *virtual machine monitor* or a *hypervisor*.

Depending on the case, the virtual environment may be based on complete software emulation of the guest hardware architecture, or the host machine may have hardware support for virtualizing itself or another (similar) architecture, which normally greatly simplifies the hypervisor design and increases performance of the virtual environment. The hypervisor implements virtual devices (video cards, disk controllers, PCI devices, etc.), which are either mapped to the real devices, device architecture permitting, or emulated in software by the hypervisor. The hypervisor may be able to manage several virtual machines simultaneously while administering the hardware resources of each. The hypervisor itself may run directly on hardware (*type 1* or *native* or *bare metal*), or under another operating system as an application (*type 2* or *hosted*).

The guest operating system itself may be totally unaware that it is running on virtualized hardware. However, more usual nowadays is that the guest system implements already some support on the level of its device drivers to optimize performance in the virtual environment. This technique is called *paravirtualization*.

Some History

The idea behind computer virtualization is old. The first ideas about a system capable of emulating another were already recorded in the late 1950’s. IBM became a major proponent of virtual machine technology in their 360, 370, and 390 series mainframes, the last two of which have been in use at CERN. Starting from the Intel 386 processor, there has been hardware support for virtualization on the PC machines. Several virtualization products have appeared since late 90s, such as VMware, Xen, KVM, Microsoft Hyper-V, etc.

USE CASE FOR ACCESS AND SAFETY SYSTEMS

Many reasons exist for using virtual machines: The original reason for IBM was to simplify OS research and design. Today’s data centres use virtual machines mainly for purposes of application isolation, load balancing, and eventually, saving of energy and money introduced by better utilization of hardware resources. However, as far as safety and access system servers are concerned, the rationale for considering server architecture based on virtual machines have to do with management of the application servers:

- Application servers can be created and configured off-line without need for the final hardware until it’s time to go live. A completely “clean” installation of an existing server can be created and dropped in with a simple switch-on/off.
- One single copy of the basic configuration can be created and all subsequent servers can be cloned from this. The only things to individually configure on each server are the network parameters and any per-server special software to be installed. It is even possible to automate all these tasks if the number of servers to be configured is large.
- Since the application servers are dissociated from the physical hardware, in case of hardware failure, the application servers can be restarted and brought back on-line within minutes, if a physical failover host is available and server back-ups are up to date.
- Being regular files on the host file system (albeit, often large), application servers can be easily backed up both locally and to a remote location.

ISBN 978-3-95450-139-7

The procedure is uniform across all servers irrespective of what applications they run.

- Virtual machine snapshot facilities of most hypervisors offer an easy way to safeguard against possible destabilizing effects of system patches. In case of problems, an old snapshot can usually be restored within minutes.
- Given a well-organized virtual machine backup and failover policy, recovering a crashed server usually shouldn’t require administrator skills in the application servers. Detailed procedures have been drafted, so that non-virtualization experts on call could easily perform these actions.

The most expensive thing when running a service is manpower. Virtualization technology has the potential of speeding up management, maintenance, and incident response times of the safety and access systems, whose smooth functioning all CERN users depend on.

The potential economic savings of not having to invest in real hardware for each individual application server are not to be discounted either: While a server machine able to run several virtual servers requires more memory, disk, and CPU power than a single application server alone, a machine able to run a dozen application servers without problems may not cost more than 50% above a normal professional-grade single server.

A new way of managing the servers introduces a new set of challenges. Virtualizing the server infrastructure generates a new abstraction layer between the actual hardware and the application servers. In a way, this shifts the complexity of server management up one level. In the end, however, the effect is positive: In a virtual environment, application server specialists are usually required only for initial configuration and functional modification of the application server itself, while managing the running, backup, and restore of the virtual servers only requires skills in that discipline, and it is the same for every server.

A new project, Virtual Safety Server Infrastructure (VSSI) was started to apply and develop virtualization for the coming and existing personnel safety and access systems at CERN.

SELECTION OF THE VIRTUALIZATION PRODUCT

The first task in the project was to investigate various leading virtualization products on the market for features and suitability against a set of requirements:

- Emphasis was on management and usability with the goal that the finalized installations be easily operable by personnel with limited expertise on virtual system management.
- It should be possible to conveniently run several virtual machines on a single server.
- It should be possible to run the hypervisor bare metal on HP server hardware as well as a client inside a host OS.
- There should be options for assuring redundancy and high availability.
- It should be possible to virtualize special hardware (PCIe and PCI-X cards, USB devices, etc.).
- The system should be compatible with dedicated SCADA and access software architectures: Siemens WinCC, PCVue, Gegelec Evolynx.
- There should be a graphical management interface.
- It should be possible to carry out automatic and manual snapshotting and backup and restore.
- It should be possible to supervise hypervisors and VMs preferably via SNMP or similar protocol.
- It should be possible to define easy procedures for maintenance team for emergency operations.
- It should be possible to run the system on different public or private networks without problems with access and maintainability.

The products originally compared were VMWare ESXi [1], Microsoft Hyper-V [2], and Xen [3], which are all in use at CERN. KVM [4] was included in the comparison later on. A first comparison was carried out on paper by comparing published features and user experiences of the products. In a nutshell, the results of the comparison were:

- Microsoft Hyper-V does not support direct pass-through of USB and PCI devices essential for running LHC and PS access servers, which interface with PCI data acquisition cards as well as USB devices (licence dongles etc.).
- Xen and KVM were tested on a test bench system. While being free software and fulfilling the technical requirements, installation and management of the hypervisors was found to require considerable time and expertise with Linux installations, module compilations, and packages, which becomes particularly challenging in a private network without connectivity to various software repositories either at CERN or outside.
- VMware ESXi 5.0 was tested with a limited duration test licence. The installation of the hypervisor on an SD card on the server motherboard was straightforward, and the system was up and running in a couple of hours.

Out of the hypervisors tested, VMware was found to best fulfil all the technical requirements, and as a supported commercial product, the look-and-feel and documentation are adequate. Also, since the licence price for a basic configuration was quite reasonable, VMware was chosen as the most suitable product for the project.

DESIGN AND IMPLEMENTATION

VSSI Architecture

The general design criteria of the system were strongly affected by the need to give highest priority to secure operation of the hosted applications. Also, the different access and safety systems needed their own isolated implementations, which meant that instead of creating a

ISBN 978-3-95450-139-7

single virtual service for all the systems, several identically configured instances would be needed. The main design criteria could then be summarized as follows:

- Each instance would consist of two identically configured physical host servers.
- The virtualized application servers (between 6 and 12 depending on the service) would be divided roughly evenly between the hosts.
- Any redundant virtual application servers would be configured to run on different hosts.
- Each host should be powerful enough to run all virtual application servers alone in case of a breakdown or maintenance of the other host.
- All virtual machines would be regularly backed up to an external disk server attached via a fast network link.
- Mass storage for data intensive applications, such as video recording, would be stored on an external disk server attached via a fast network link.

A schematic of the redundant dual host architecture is shown in Figure 1.

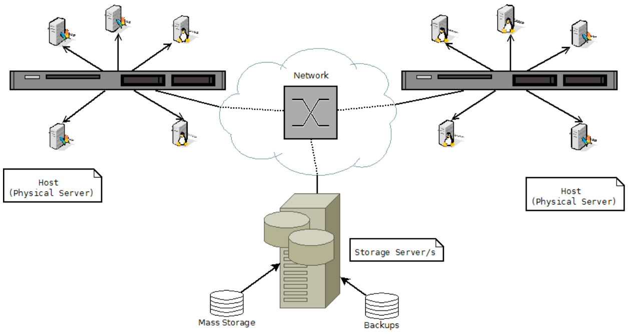


Figure 1: VSSI redundant dual host architecture. Virtual machines are divided between two physical hosts with a separate disk server for external storage. All hosts are connected to a local dedicated network switch.

Hardware and Software

New server hardware was acquired for the VSSI project. As the access team has a long history of HP servers and their management, the chosen models were HP Proliant DL360 G7 with dual Intel Xeon X5675 CPUs, 32GB RAM, six 146GB 6G SAS 15kRPM disks configured in RAID 6, and dual power units for maximum reliability [5].

The hypervisor was VMware ESXi 5.0.0 with VMware Data Protection for backup of the virtual machines on an external NFS share. HP Integrated Lights Out (ILO) management interface was configured for out-of-band low-level access to the servers [6].

Application to Access and Safety Systems

VSSI is currently in use by four access and safety systems at CERN:

- **LACS** (LHC Access Control System) controls who enters the LHC and when. All LACS test and development servers have been virtualized at the LACS test bench. LACS is based on Evolynx software running on Windows Server 2003, and the full access control system comprises several application servers:
  - A master server that controls all the other servers. Direct access to the RS232 port to communicate with external hardware.
  - "Frontal" servers that communicate with on-site access controllers and PCs.
  - A biometric server that manages the iris-codes and communicates with on-site authentication controllers.
  - Two special frontals that communicate with the gateway PLC to the LHC Access Safety System (LASS) residing in its own isolated private network. Direct access to an Applicom data acquisition card on the PCI-X bus.

  Experience gained on the test platform is quite satisfying as far as management and performance are concerned, which promises well for the future update of LACS production servers.
- **PACS** (PS Access Control System) controls who enters the PS complex and when. The new PS access control system, which is being installed in 2013-2014, uses virtual servers from the start [7]. Two identical host machines are installed. The main application servers running on those hosts are:
  - Two redundant access control servers running Siemens WinCC on Windows 2003. Direct access to a Siemens data acquisition card CP1623 on the PCIe bus.
  - A WinCC webserver to serve synoptic displays to panel-PCs on sites.
  - Two video servers that run video software on Windows 2008 and store video archives on an external disk server.
  - An access control server that runs access control software managing zone authorizations.
  - A monitoring proxy server for Safety System Monitoring (SSM) [8].

  There is also a test bench system, which contains an identical set of servers to the production system.
- **CSAM** (CERN Safety Alarm Monitoring) manages alarms for the fire brigade. Five virtual servers run instances of PCVue SCADA software each managing various parts of the CSAM system.
- **SSA** (Safety System Atlas) is the specific personnel safety system for the ATLAS detector. There is one host, with one virtual server to run a WinCC application. Another WinCC server is installed as a cold spare.

Production virtual machines are backed up regularly on an external disk server. The backup process takes snapshots automatically and transparently of running virtual machines and stores them on disk. However, virtual machines with direct access to PCI devices cannot be backed up while running because the hypervisor cannot snapshot an external device not directly under its

ISBN 978-3-95450-139-7

control. A manual intervention must be scheduled from time to time to back up those servers. In the future it may be possible to automate this process as well by scripting the hypervisor. Figure 2 presents a schematic of the systems and hosts in the VSSI framework.

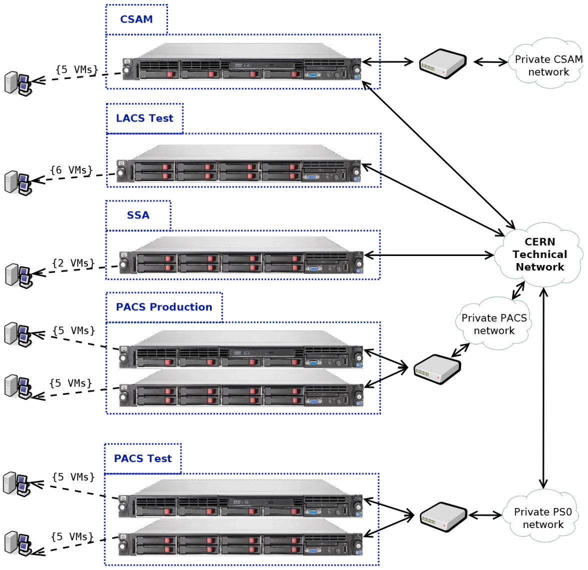
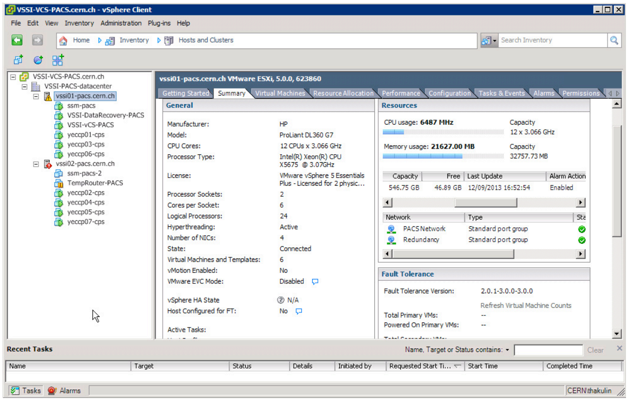


Figure 2: Schematic of the VSSI framework today.

Virtual Machine Management Interface

Virtual machines are managed by the hypervisor, whose job is really just to act as an arbitrator of hardware resources between the virtual machines and to allow accessing and managing them in a controlled manner. The basic VMware hypervisors include just a simple command line to do these tasks. A commercial VMware vCenter licence buys a graphical management interface, which can be used to access individual hypervisors directly or via a separate vCenter server (itself running in a virtual machine) able to manage large clusters of hypervisors. The basic VMware management interface of PACS virtual servers is shown in Figure 3.



General

Manufacturer: HP
Model: ProLiant DL360 G7
CPU Cores: 12 CPUs x 3.066 GHz
Processor Type: Intel(R) Xeon(R) CPU X5675 @ 3.07GHz
License: VMware vSphere 5 Essentials Plus - Licensed for 2 physic...
Processor Sockets: 2
Cores per Socket: 6
Logical Processors: 24
Hyperthreading: Active
Number of NICs: 4
State: Connected
Virtual Machines and Templates: 6
vMotion Enabled: No
VMware EVC Mode: Disabled
vSphere HA State N/A
Host Configured for FT: No
Active Tasks:

Resources

CPU usage: 6487 MHz Capacity 12 x 3.066 GHz
Memory usage: 21627.00 MB Capacity 32757.73 MB

Capacity	Free	Last Update	Alarm Action
546.75 GB	46.89 GB	12/09/2013 16:52:54	Enabled

Network	Type	Sta
PACS Network	Standard port group	✓
Redundancy	Standard port group	✓

Fault Tolerance

Fault Tolerance Version: 2.0.1-3.0.0-3.0.0
Refresh Virtual Machine Counts
Total Primary VMs: --
Powered On Primary VMs: --

Name	Target	Status	Details	Initiated by	Requested Start Ti...	Start Time	Completed Time

Figure 3: VMware hypervisor management interface. On the left is the list of virtual machines running on two separate hosts. On the right is the summary page of the host number one with configuration and status info.

VMware hypervisor also collects statistics of the hosts as well as all the virtual machines, which can be viewed as historical or real time graphs as presented in Figure 4.

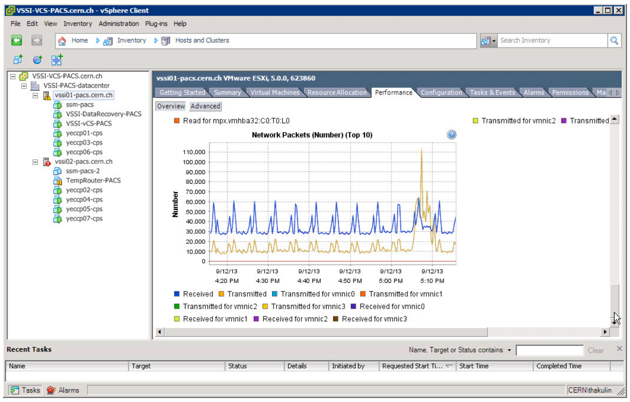


Figure 4: VMware host performance graph showing network packet statistics of all host network interfaces.

Access Teams as Users and Maintainers

One integral part of the VSSI project was to create a system, which would be manageable by access and safety system maintenance teams, who are no particular experts in virtual infrastructure management. For that purpose, three user roles have been defined on the hypervisor: *Operator* who manages the application servers without need to reconfigure the virtual machine parameters, *Power Operator* who is able to change virtual machine parameters (memory, CPU, network interfaces, etc.), and *Administrator* who can do anything. For each role there is a corresponding set of user documentation.

CONCLUSIONS

Virtual Safety System Infrastructure (VSSI) project was launched to introduce modern virtual server management techniques to managing CERN access and safety systems. The new PS access control system, CSAM servers, and the test servers of the LHC access control system were successfully virtualized, and corresponding operation and management procedures defined. Near term plans include moving the old LACS production servers to VSSI, and most importantly, gaining experience and feedback from the maintenance and operation teams and first-line on-call interventions.

REFERENCES

[1] http://www.vmware.com
[2] http://www.microsoft.com/hyper-v-server
[3] http://xen.org
[4] http://www.linux-kvm.org
[5] http://www8.hp.com/us/en/products/proliant-servers
[6] http://en.wikipedia.org/wiki/HP_Integrated_Lights-Out
[7] P. Ninin et al., "Refurbishing of the CERN PS Complex Personnel Protection System," MOPPC059, these proceedings.
[8] T. Hakulinen et al., "Revisiting CERN Safety System Monitoring (SSM)," MOPPC055, these proceedings.