EXPERIENCE OF VIRTUAL MACHINES IN J-PARC MR CONTROL

N. Kamikubota[#], S. Yamada, K.C. Sato and N. Yamamoto, J-PARC, KEK & JAEA, Ibaraki, Japan
T. Iitsuka, S. Motohashi, D. Takahashi and S. Yoshida, Kanto Information Service, Ibaraki, Japan
H. Nemoto, ACMOS Inc., Ibaraki, Japan

Abstract

In the J-PARC Main Ring, we have used the virtual technology extensively in our accelerator control. In 2011, we developed a virtual-IOC, an In/Out Controller running on a virtual machine. Several virtual IOC's were introduced in the accelerator operation. In 2012, we ported control services (dhcp, ldap, rdb, achiver, etc.) into virtual machines. Implementation details and experiences are reported.

INTRODUCTION

Recently, the virtualization technology has been developed and spread remarkably. In the past, we needed technical skills and an appropriate budget to use it. Now non-expert can use the technology. Recent Linux and Windows versions include a virtual machine monitor: in the case of Linux, KVM (Kernel-based Virtual Monitor), was merged into Linux kernel after 2.6.20. This is because broad users have needed backward-compatibility of 32-bit systems using modern 64-bit PC platform.

When we introduce the virtualization technology, structure changes take place in a server computer system. Before 2010 without the technology, in order to simplify server maintenance, one server computer handles one service. As a result, we have many pieces of low/middle-spec computers (Figure 1, left). Blade-type server systems had been available in the commercial market, and utilized for such server systems. Recent years with the technology, a few high-spec computers are used as host machines. One host handles many virtual machines. However each virtual machine handles one service, just as is before, to keep simple maintenance ability (Figure 1, right).


Figure 1: Structure changes in a server computer system.

In the virtual environment, we can: (a) move a virtual machine from a host to another, and (b) change hardware spec of a virtual machine easily. These features enable flexible re-arrangement of virtual machines over host

_____
[#]norihiko.kamikubota@kek.jp

machines. In addition, one can expect cost reduction in total.

PAST STUDIES IN J-PARC MR

The J-PARC Main Ring (hereafter MR) is a high-power proton synchrotron with beam-energy 30GeV. The control system has been developed using the EPICS (Experimental Physics and Industrial Control System) toolkit [1]. The control system has been used for the beam operation of MR since 2008 [2].

In the control system, more than 100 IOC's (In/Out Controller) are used [3]. Most of them are: (a) VME-bus computers, a traditional platform in this field, and (b) PLC-based controllers with EPICS embedded [4].

Around 2010, we carried out feasibility studies on virtualization technologies such as VMware and XEN. We developed a virtual-IOC (hereafter "vioc"), an IOC running on a virtual machine. In 2011, part of VME-bus IOC's with no real I./O (so-called soft IOC), or those for networked devices, were shifted to vioc's. The detail descriptions are given in [5]. In 2013, about 30 vioc's are introduced.

The success of vioc encouraged us to extend the virtual environment more in our control. In 2012, we ported control services, such as dhcp, ldap, rdb, channel achivers, etc., into multiple virtual machines. They have been used in MR operation since then.

VIRTUALIZATION EXPERIENCES

Virtual Machines Configurations

In the J-PARC MR control system, we have adopted Scientific Linux (hereafter SL) as the main OS (Operating System). In 2008, we started with SL4. During the summer of 2012, we upgraded the main OS to SL6, which has native support of KVM. Thus, we decided to use KVM as a default virtual monitor.

In 2013, we have two groups of virtual machines: one for control services, and another for vioc servers. Both groups use SL6.0 as a host OS. Each group consists of three host computers. The host is a high-spec blade-type computer (IBM Blade HS22, Xeon E5504 (4core 2cpu) 2GHz) with 20 GB of memory. In each group, we distribute virtual machines to keep a good load balance between three hosts. When we have a host trouble, we stop the host and keep to operate the group using other two hosts. When we need more CPU power and/or memory allocation in the future, we will add more host computer(s) to the group.

The configuration for the control services group is shown in Figure 2. Eight server computers are virtualized, and summarized into three host machines. Typical
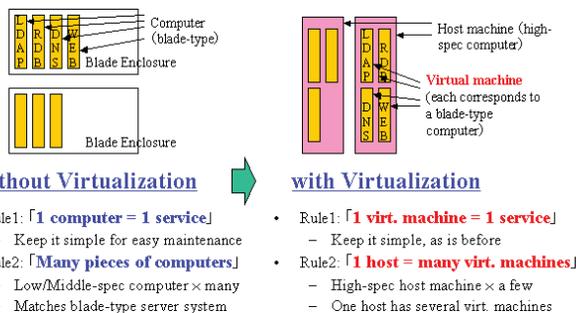
memory allocation to each virtual machine is 2-4GB. The configuration for the vioc servers group is shown in Figure 3. About 30 virtual machines run on three host machines. The 23 of them are used in daily MR operation, and the rests are for individual studies. Typical memory allocation to each virtual machine is 512MB. One host (jkjblade3a) belongs to the both groups temporary, thus, total number of hosts is five.

The virtual machine "simulator server", assigned in the control services group, has a unique functionality. It keeps an old development environment of SL4.4. It is used to check backward-compatibility of control applications between SL4 and SL6.

Host machine	Virtual Machine
SL6.0 memory=20GB	SL5.4 memory=2-4GB
(host1) jkjblade**3a**	· channel archive engines · channel access gateway
(host2) jkjblade**3b**	· admin. server (dhcp, tftp, ldap slave) · application server (cron, zlog, cacti) · RDB server (postgres, mysql)
(host3) jkjblade**3c**	· admin. server (ldap master) · simulator server (old OS: SL4.4)

Figure 2: Host and virtual machines for control services.

Host machine	Virtual Machine
SL6.0 memory=20GB	SL6.3 memory=512MB
(host1) jkjblade**3a**	· cont-group: 3 vioc
(host2) jkjblade**3e**	· mag-group: 4 vioc · inj-group: 2 vioc · mon-group: 3 vioc · accom-group: 1 vioc · rf-group: 1 vioc
(host3) jkjblade**3f**	· cont-group: 6 vioc · sx-group: 2 vioc · mon-group: 1 vioc · individual study: 6 vioc

Figure 3: Host and virtual machines for virtual IOC's.

Operation and Management

The system status of host machines can be checked using standard tools. In Figure 4, the 10-minute trends of CPU load and memory usage are shown for typical host machines (jkjblade3b, jkjblade3e). The memory usage of jkjbalde3b is as high as ~18GB, which corresponds to 90 % of the total memory. In the future, adding another host computer is preferable to the control services group.

A standard tool, virtmgr, has been used for various management purposes. When we need to move a virtual machine from a host to another, this tool is used. Figure 5 (left) is a screenshot of re-locating procedure by the tool. In addition, a web-based management screen was developed to show an overview of host and virtual machines. It indicates host machines together with corresponding virtual machines in a list format (Figure 5, right).

Figure 4: CPU load and memory usage of typical hosts.
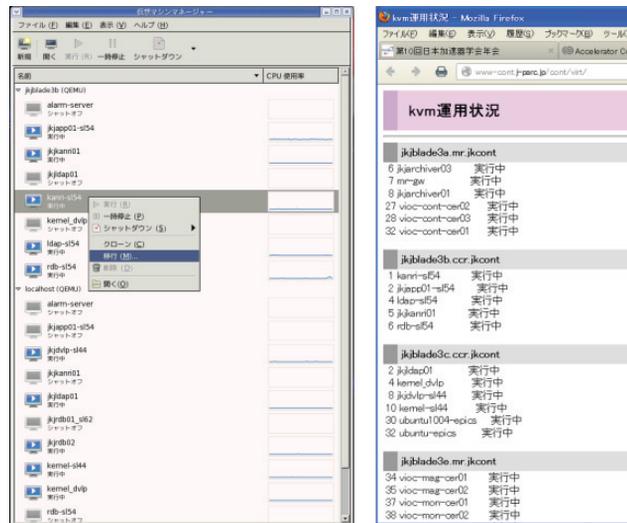
Figure 5: Management tools for virtual machines.

Trouble - Unexpected Behaviour of NFS

When we started operation for the control services in 2012, we used SL6.0 for virtual machines. We often suffered unexpected behaviours of virtual machines that NFS-mount turns to read-only, followed by a network stop by maintenance and/or fault. In such cases, manual reboot is required for each virtual machine. The behaviour never occurs with non-virtual SL6 computers.

The SL6 adopted NFS v4 instead of v3. Thus, we surveyed NFS parameters suitable for a virtual machine, but not succeeded yet. The behaviour occurs only when a virtual machine is SL6. At present, we decided to use SL5.4 for the virtual machines of control services, in order to avoid this problem (Figure 2).

Trouble – Sudden Stop of a Host Machine

On February 9, 2013, a host computer (jkjblade3f) stopped suddenly. Accordingly many vioc's were stopped

Software Technology Evolution

during MR operation. The stopped vioc's were moved to another host (jkjblade3e), and restarted manually. The recovery process took 3 hours.

The later analysis of the system log showed that a CPU request from a virtual machine to the host had been suspended over a limit. However, causing a sudden stop of host seems a known bug. We re-configured the Linux kernel parameter, soft_lockup, from 10 to 60. After that, no stop has been happen.

If a host is alive, re-location of virtual machines is easy by a tool. However when a host is stopped unexpectedly, we are requested to start virtual machines manually. Automatic recovery procedure must be discussed in the future.

CONCLUSION

In J-PARC MR, a virtual environment using KVM has been used in the accelerator operation. We started to use virtual-IOC's, EPICS IOC's running on virtual machines, since 2011. Virtual machines for control services started in 2012. Although we experienced some troubles, the overall operation of virtual machines has been successful.

REFERENCES

[1] http://www.aps.anl.gov/epics/

[2] N. Kamikubota et al., "J-PARC Control toward Future Reliable Operation," ICALEPCS 2011, Grenoble, France, Oct. 2011, MOPMS026, pp. 378-381; www.JACoW.org

[3] H. Nemoto et al., "IOC Surveillance System for J-PARC MR Control" (in Japanese), PASJ Annual Meeting in Osaka, Japan, Aug. 2012, pp. 754-747.

[4] J.-I. Odagiri et al., "Application of EPICS on F3RP61 to Accelerator Control," ICALEPCS 2009, Kobe, Japan, Oct. 2009, THD005, pp. 916-918; www.JACoW.org

[5] N. Kamikubota et al., "Virtual IO Controllers at J-PARC MR Using XEN," ICALEPCS 2011, Grenoble, France, Oct. 2011, WEPMU039, pp. 1165-1167; www.JACoW.org