Experience of Virtual Machines in J-PARC MR Control

N.Kamikubota[A]   S.Yoshida[B]   S.Yamada[A]   K.C.Sato[A]   S.Motohashi[B]   T.Iitsuka[B]   H.Nemoto[C]   D.Takahashi[B]   N.Yamamoto[A]

A) J-PARC / KEK (Tokai Campus)   B) Kanto Information Service (KIS)   C) ACMOS Inc.

Abstract

At the J-PARC Main Ring (MR), we have used virtual-machine environment extensively in our accelerator control.
In the summer of 2012, we updated our operating system from Scientific Linux 4 (SL4) to Scientific Linux 6 (SL6). In the SL6, KVM virtual-machine environment is supported as a default service. This fact encouraged us to port basic control services (dhcp, tftp, ldap, rdb, achiver, etc.) to virtual machines.
Virtual machines are running on a few (not many) physical machines. This scheme enables easier maintenance of control services than before.
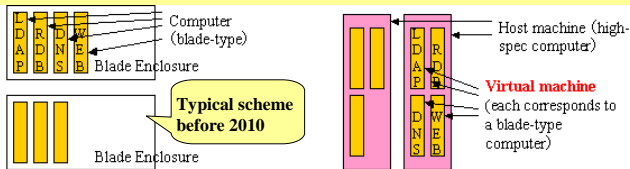
Using Virtualization Technology

- **Virtualization Technology in Recent Years**
  - **64bit-OS supports virtual machine monitor as default**
    - Virtual PC on Windows   test compatibility with a 32-bit Windows XP
    - KVM on Linux i.e. RHEL 6, Scientific Linux 6, ..

- **Virtualization at Server-side**
  - Virtualization has changed server structures

without Virtualization

- Rule1: 「1 computer = 1 service」
  - Keep it simple for easy maintenance
- Rule2: 「Many pieces of computers」
  - Low/Middle-spec computer × many
  - Matches blade-type server system

with Virtualization

- Rule1: 「1 virt. machine = 1 service」
  - Keep it simple, as is before
- Rule2: 「1 host = many virt. machines」
  - High-spec host machine × a few
  - One host has several virt. machines

One can move a virtual machine from a host to another,
depending on host machines load status, and on maintenance schedule, and so on ..
Save computer resources and maintenance cost in total

Past Studies at J-PARC MR

- **Virtual IOC (2011)**
  - **EPICS IOC (I/O controller) running on a virtual machine**
    - Compatible to an IOC running on a "real" VME-bus computer
    - For soft-records and simple network devices (no real I/O)
    - 2010: Feasibility studies using XEN around 2010
    - 2011: Operational version using KVM started
    - Now (2013): about 30 virtual IOCs for daily operation of J-PARC MR

IOC – real & traditional   IOC – virtual => "vioc"
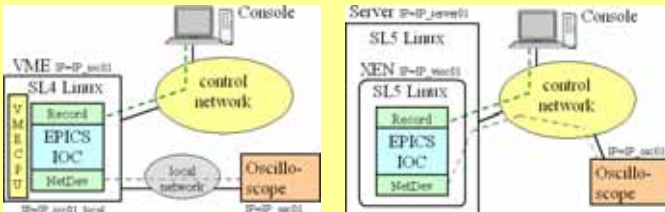
- **References:**
- J-PARC control in general
  - J-PARC CONTROL TOWARD FUTURE RELIABLE OPERATION , ICALEPCS2011, N.Kamikubota et al.
  - IMPROVEMENT OF COMPUTER SYSTEMS FOR J-PARC MR CONTROL,   (2012) in Japanese, N.Kamikubota, et al
- EPICS IOC and Virtual IOC at J-PARC MR
  - VIRTUAL IO CONTROLLERS AT J-PARC MR USING XEN , ICALEPCS2011, N.Kamikubota et al.
  - IOC Surveillance System for J-PARC MR Control   (2012), in Japanese, H.Nemoto et,al.
  - Environment of Linux-based IOC for J-PARC MR, PCaPAC2006, N.Kamikubota et al.

Virtualization Setup and Experience at J-PARC MR

- **Virtual Servers for Control Services**
  - **Basic control services are moved to virtual machines** (2012)
    - dhcp, tftp, ldap, RDB (postgres, mySQL)
    - cron, zlog (E-log), cacti, channel archive engines
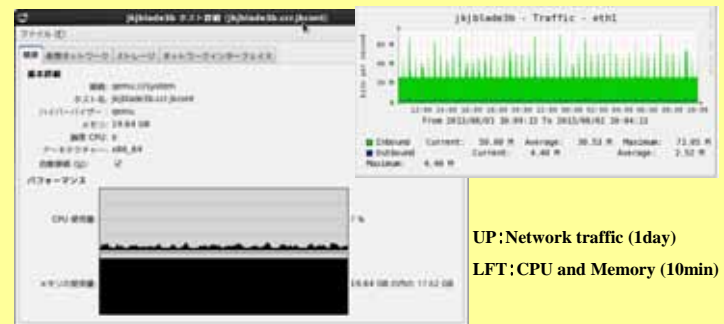   - Simulator machine for an old linux version

- **Virtual Environment Setups** in 2013
  - **2 groups – service servers and vioc servers**
    - Each group consists of 3 host machines
    - Each group can be in operation with 2 hosts (in case one host dead)
    - In each group, distribute virtual machines to have a balance between 3 hosts

Service Servers

Host machine	Virtual Machine
SL6.0 memory=20GB	SL5.4 memory=2-4GB
jkjblade3a	·Channel archive engines ·Channel access gateway
jkjblade3b	·admin. server (dhcp, tftp, ldap slave) ·Appl. server (cron, zlog, cacti) ·RDB server (postgres, mysql)
Jkjblade3c	·admin. server (ldap master) ·simulator server (old OS: SL4.4)

VIOC Servers

Host machine	Virtual Machine
SL6.0 memory=20GB	SL6.3 memory=512MB
jkjblade3a	·cont-group: 3 vioc
jkjblade3e	·mag-group: 4 vioc · inj-group: 2 vioc ·mon-group: 3 vioc ·accom-group: 1 vioc ·rf-group: 1 vioc
Jkjblade3f	·cont-group: 6 vioc ·sx-group: 2 vioc ·mon-group: 1 vioc · for test: 6 vioc

- **System load of a typical host machine** (ex. jkjblade3b)

UP Network traffic (1day)

LFT CPU and Memory (10min)

- **Experienced Troubles**
  - **a) Sudden stop of a host machine**

2013/2/09 20:19 a host, jkjblade3f, stopped
  => vioc's were moved to another host (3e) manually
      23:40 recovery procedure was completed
  => **soft_lockup** (kernel parameter) **- setup changed**
      "kernel: BUG: soft lockup - CPU#4 stuck for 17163091968s! [qemu-kvm:20252] "
      known bug: virt. OS dies when cpu not assigned from a host

  - **b) NFS of a virtual machine is unstable**

NFS mount of a virtual machine turns to read-only unexpectedly,
      after network maintenance and/or faults
  => It happens only when a virtual machine is SL6.
      Virtual machines for service servers were down-ported to
      SL5.4, to avoid this problem.

Problems:
When a host machine stopped, recovery procedure is not automatic.
proper NFS configuration for SL6 virtual machines must be studied more.