

Lecture 6: February 8

*Lecturer: Prashant Shenoy**Scribe: Harihara Subramanian Krishnamoorthy*

6.1 Introduction

An example of hardware virtualization is VirtualBox. An example of OS Virtualization is Docker

6.2 Memory Virtualization

- The OS manages the pagetables and tracks pages required.
- Hypervisor can modify pages in kernel mode only
- The OS runs in user mode usually.

Modifying Read-only pages causes traps to hypervisor. Hypervisor maps to actual page in shadow page table present in the OS.

6.3 I/O Virtualization

- The guest OS does read and write operations to virtual disk.
- When a virtual machine is created the hypervisor creates a blank page that is a disk for the guest OS.

6.4 NIC Virtualization

The IP assigned to the virtual machine is independent of the underlying machine. It has its own logical interface to emulate ethernet cards. It uses the underlying OS's ethernet card.

6.5 OS level Virtualization

- They run applications inside containers that is isolated from the underlying machine. No full Operating system.
- They rely on the underlying operating system and only the underlying system schedules processes.
- Processes may not be able to see the underlying operating systems.

In Solaris the containers are called zones.

6.5.1 Benefits

Isolation among processes

Bare-metal performance since its run on OS so no slowdown.

6.5.2 Disadvantage

All the processes use the same kernel so a kernel bug crashes all containers

6.6 OS Mechanisms for LXC

OS mechanisms for resource isolation and management

- namespaces: process-based resource isolation - Gives process based isolation.
- Cgroups: limits, prioritization, accounting, control - specify the memory ram / number of cores, i.e the resource limit to be allocated.
- chroot: apparent root directory - Change the root directory so that only that part of the subtree can be seen.
- Linux security module, access control
- Tools (e.g., docker) for easy management

6.7 Linux Namespaces

Namespace: restrict what can a container see?

Provide process level isolation of global resources

- Processes have illusion they are the only processes in the system
- MNT: mount points, file systems (what files, dir are visible)
- PID: what other processes are visible
- NET: NICs, routing
- Users: what uid, gid are visible
- chroot: change root directory

6.8 Linux cgroups

Shared based: Containers are given weights. There are a whole class of algorithms. Depending upon the weight, the proportional shares are given in a deterministic fashion. The same mechanism can be used for any resources here.

So if two processes has 1 and 2 as weights they will receive 1/3 and 2/3 of the CPU time.

Hard limits: Hard limit means there is a hard allocation, no matter if a process uses it or not, it is allocated.

6.9 Docker

Docker uses linux containers.

Allows portable container deployment.

Makes software deployment easier because its self contained as an image. Sometimes OS is also packages VM ware supports this.

Docker uses union file system creates only the difference

6.10 Uses of Virtualization

Data Centers

reduces cost due to multiple virtual servers packed together

Cloud Computing

allows mapping of virtual servers to host

Desktop Computing

Can do multiplatform development

6.11 PlanetLab

Virtualized architecture used for research by students in different locations