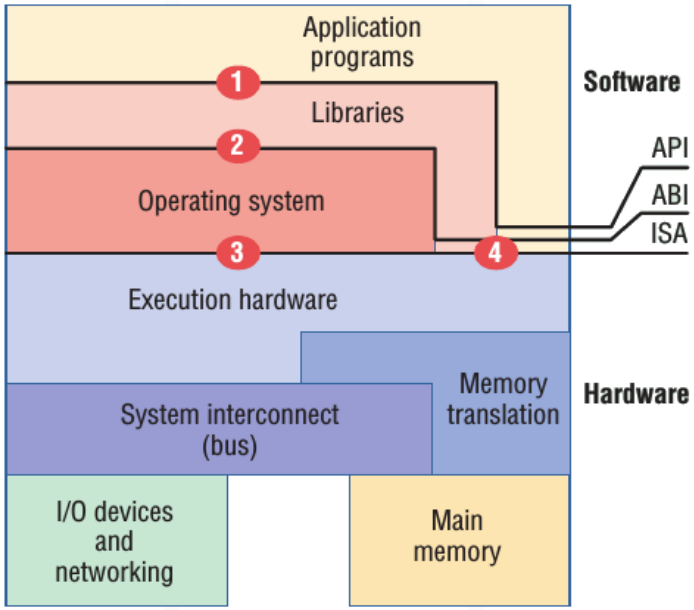
Màquines Virtuals i Virtualització del SO

Sistemes Operatius II

lluis.garrido@ub.edu

Grau d'Enginyeria Informàtica

Un sistema operatiu, una màquina virtual?



Un sistema operatiu es pot veure com una màquina estesa:

- Oculta a l'usuari de tots els detalls escabrosos que han de ser realitzats per accedir als dispositius.
- Ofereix a l'usuari una màquina virtual i una una interfície (les crides a sistema), molt més senzilla d'utilitzar.

Un sistema operatiu, una màquina virtual?

El sistema operatiu permet l'execució de processos:

- Cada procés té la il·lusió de tenir la màquina completa per sí sola.
- Cada procés pot executar instruccions directament sobre la màquina en mode usuari. A més, cada procés pot accedir als dispositius d'entrada-sortida a través del sistema operatiu.
- Un error de programació o malícia en una aplicació no afecta a la resta de processos.
- Un procés no hauria de poder consumir tots els recursos (CPU, memòria, ...) en detriment d'altres processos.

El sistema operatiu aconsegueix tot això gestionant el recursos disponibles així com amb tècniques de temps compartit (*time-sharing*) del maquinari.

Un sistema operatiu, una màquina virtual?

De forma estricta, un sistema operatiu no es per definició una màquina virtual. Què és un sistema operatiu?

- És un programari que té accés directe al maquinari i ofereix un nivell d'abstracció respecte els processos que s'hi executen.
- Està dissenyat per aprofitar l'arquitectura del maquinari per assegurar la seguretat en executar processos, donar la il·lusió que hi ha memòria "infinita", etc.

Té restriccions...

- Ha estat dissenyat per a una determinada arquitectura.
- S'executa per a un determinat conjunt instruccions (Intel PC, PowerPC, ...)
- Se suposa que és l'únic programari que controla el maquinari.

Limitem doncs la flexibilitat del sistema.

Què és una màquina virtual? Objectius

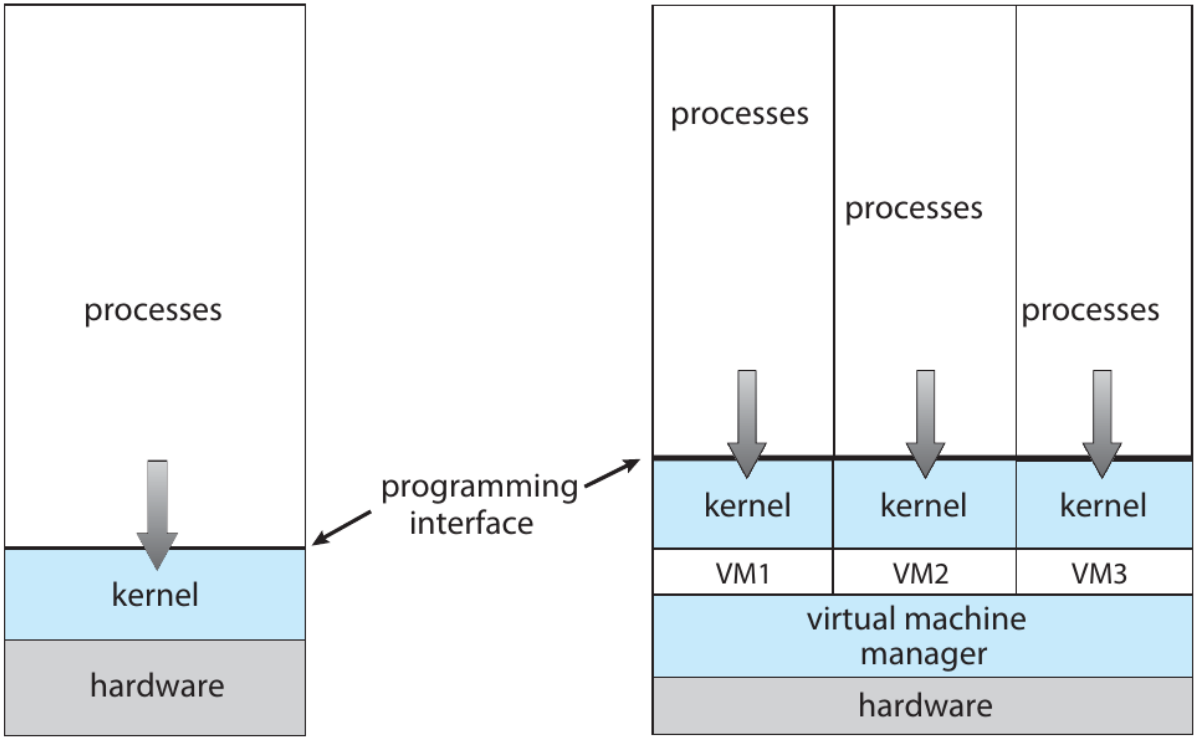
Què és, aleshores, la virtualització?

- La virtualització té com a objectiu donar més flexibilitat a les restriccions anteriors. Quan es virtualitza un sistema (processador, memòria, dispositius d'entrada-sortida, ...) la seva interfície i recursos es mapen a la interfície i recursos del sistema real. En conseqüència, el sistema real es transforma de manera que sembla ser un sistema virtual diferent[1].

- Segons la Wikipedia, "una màquina virtual (VM) és la virtualització/emulació d'un sistema informàtic. Les màquines virtuals es basen en arquitectures d'ordinadors i proporcionen la funcionalitat d'un ordinador físic. Les seves implementacions poden implicar maquinari especialitzat, programari o una combinació dels dos."

[1] Una tarja gràfica o de xarxa es pot mapar a una tarja gràfica o de xarxa diferent al sistema virtual.

Què és una maquina virtual?



Les màquines virtuals "habituals" involucren diverses components

- Un Virtual Machine Manager (també anomenat hypervisor)
- Diverses màquines virtuals (Virtual Machines)

Què és una maquina virtual?

Les màquines virtuals involucren diverses components

- El host és correspon al maquinari físic.
- La Virtual Machine Manager (també anomenada hypervisor) crea i permet gestionar màquines virtuals (Virtual Machines, VM), proveint a cada VM una interfície exacta al host.
- La màquina virtual (VM) és la que virtualitza/emula i proveeix la funcionalitat d'un ordinador físic (pot oferir un maquinari diferent del que hi ha físicament).
- Es poden executar diverses màquines virtuals a la VMM, i proveeixen al procés guest d'una còpia virtual del host. Típicament el procés guest és un sistema operatiu.

Alguns detalls tècnics...

Cada sistema operatiu es dissenya per tenir un control total del sistema (el maquinari). A una màquina virtual el VMM...

- És qui fa creure a cada sistema operatiu que té control total del sistema (ara s'executa en mode usuari!).
- Els nous processadors inclouen instruccions màquina que "faciliten" aquesta tasca: a l'arquitectura x86 es coneixen com a instruccions VTX.

Les instruccions VTX permeten fer l'anomenat "trap-and-emulate":

1. Un sistema operatiu guest fa una operació "no permesa" (pex escriure dades a disc, dibuixar a pantalla, ...).
2. El processador captura ("trap") aquestes instruccions.
3. El VMM executa el que vol fer el sistema operatiu guest. Gràcies al mapat realitzat per la màquina virtual s'executa allò que el sistema operatiu guest volia fer.

Tipus de màquines virtuals

Els type 1 hypervisors es poden trobar als servidors grans

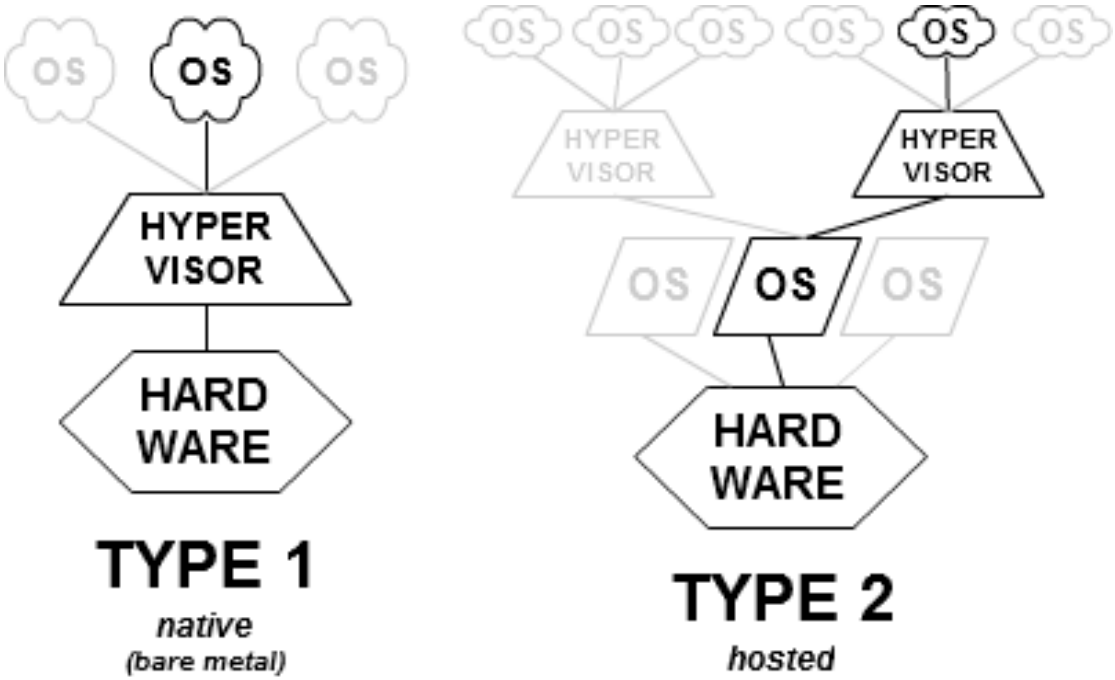
- La VMM s'instal·la a l'ordinador. En engegar l'ordinador, aquest s'executa directament sobre el maquinari i proveeix serveis tradicionals com la planificació o la gestió de memòria.
- La VMM és, de fet, un sistema operatiu que permet executar i gestionar altres sistemes operatius. Existeixen múltiples VMM d'aquest tipus, com per exemple el Citrix XenServer o el VMWare ESX[2].
- Una única màquina pot doncs executar múltiples sistemes operatius a la vegada, cadascuna a la seva pròpia màquina virtual[3]. Cada sistema operatiu creu que té control total sobre la màquina (virtual). La virtualització és doncs un tècnica per proveir a cada guest d'un maquinari (virtual)...

[2] No el confoneu amb el VMWare Workstation que utilitzeu habitualment al vostre ordinador

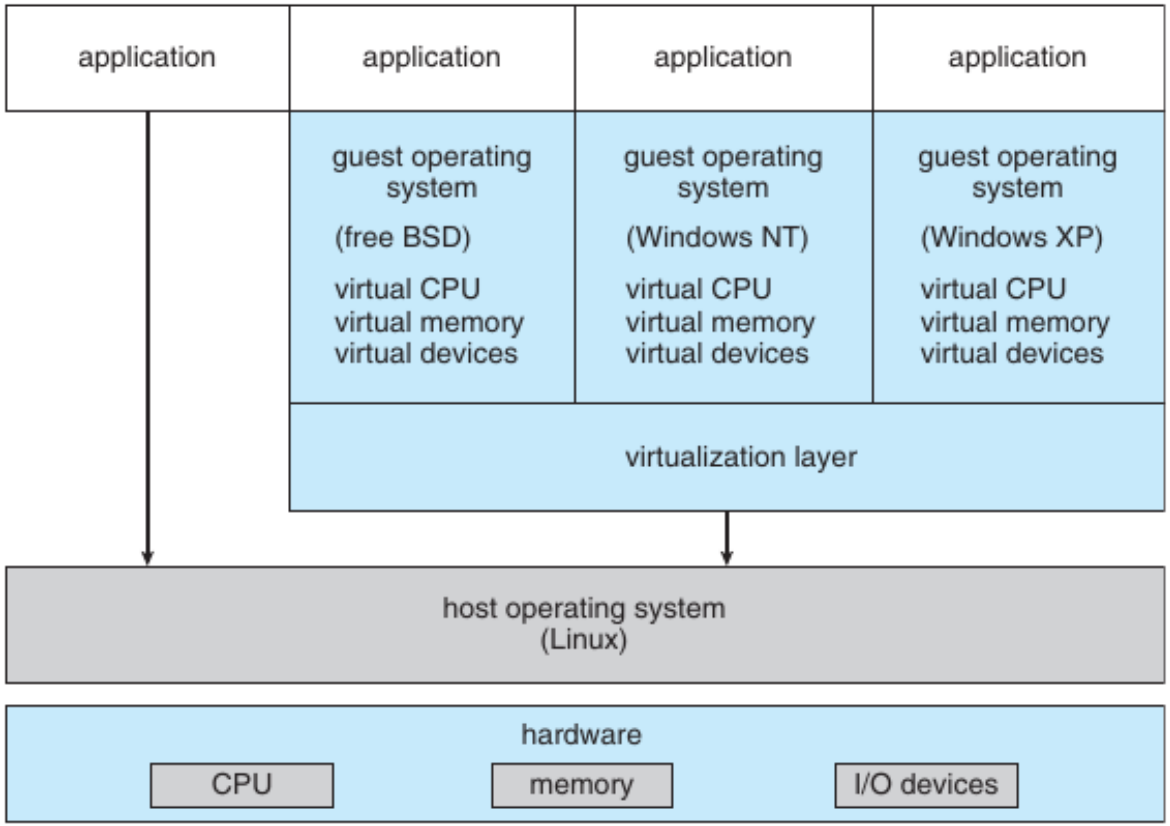
[3] Per exemple, la màquina pot tenir 4 processadors i es pot assignar una màquina virtual a cada processador.

Tipus de màquines virtuals

Els type 2 hypervisors s'executen sobre un sistema operatiu convencional. El sistema operatiu guest s'executa com si fos un procés al host



Exemple: VMware Workstation



Type 2 hypervisor: WMware Workstation executa a sobre de Windows i Linux i permet executar diversos sistemes operatius, cadascú a la seva màquina virtual.

Exemple: VMware Workstation

Sobre el WMware Workstation

- El WMWare té la seva capa de virtualització que abstreu el maquinari en màquines virtuals independents que permeten executar sistemes operatius. Cada màquina virtual té la seva pròpia CPU virtual, memòria, discos, xarxa, etc.
- El disc d'una màquina virtual no és més un fitxer al host. En copiar el fitxer podem crear un guest idèntic a l'original.

Tipus de màquines virtuals

Les màquines virtuals es poden classificar també segons la seva funcionalitat

- Màquines virtuals de sistema: proveeixen un substitut d'una màquina virtual i permeten executar un sistema operatiu. Inclouen els type 1 and type 2 hyptervisors.
- Màquines virtuals de procés: permeten executar un procés en un entorn independent de la plataforma. Les VMM no virtualitzen el maquinari real sinó un sistema virtual optimitzat. Un exemple és el Java. En compilar el codi font es generen bytecodes, un "codi màquina abstracte", que interpreta i executa la màquina virtual de Java.

Beneficis i característiques

Les màquines virtuals

- Van aparèixer cap al 1972, als *mainframes* de IBM, fent servir el IBM VM, el qual proveïa les màquines virtuals necessàries. Aquest sistema ha evolucionat i encara s'utilitza avui en dia.
- Amb l'aparició dels ordinadors de sobretaula l'interès en les màquines virtuals va disminuir.
- Avui en dia les màquines virtuals tornen a tenir gran popularitat, en especial als servidors o granges de servidors que s'utilitzen a tot el món.

Beneficis i característiques

Alguns beneficis i característiques

- Virtualització d'escriptoris: una empresa pot proporcionar escriptoris virtuals amb màquines de la pròpia empresa o d'una empresa externa.
- Virtualització de servidors: una màquina física pot gestionar múltiples servidors, cadascuna a la seva pròpia màquina virtual i sistema operatiu. Si un servidor falla, els altres no queden afectats.
- Recursos: una empresa que necessiti recursos computacionals pot contractar els serveis d'una empresa externa especialitzada en gestionar la infraestructura associada a màquines virtuals.

Beneficis i característiques

Més beneficis i característiques

- Les màquines virtuals estan (pràcticament) aïllades entre sí. A l'usuari se li pot oferir una màquina virtual amb recursos "restringits" (menys CPU, menys RAM) de la realment disponible.
- La virtualització permet "aturar" una màquina virtual (amb el seu sistema operatiu), fer-ne una còpia, moure-la a una altra màquina virtual, emmagatzemar-la, etc.
- La virtualització facilita el desenvolupament i testeig de nous sistemes operatius, així com el desenvolupament i testeig d'aplicacions en diferents sistemes operatius al mateix temps.

Qui utilitza màquines virtuals?

Moltes empreses (petites i grans) utilitzen serveis externs per fer servir màquines virtuals. Les empreses utilitzen aquests serveis ja que ofereixen una gran escalabilitat (llistat any 2021):

- Google Cloud Platform: Spotify, HSBC, Facebook, Domino, Sony Music, Ubisoft.
- Amazon Web Services: Siemens, Apple, Adobe, Pfizer, Nasa, Netflix.
- IBM Cloud: Koopman Logistics, Crédit Mutuel, Osram AH, RS Components.
- Microsoft Azure: Coca-Cola, Accenture, Adobe, Citrix, Symantec Corporation.

Resum

La virtualització és un mètode per proveir d'un duplicat del maquinari subjacent

- A les màquines virtuals de sistema cada VM es pot executar un sistema operatiu diferent i cada VM pot oferir un maquinari diferent a cada sistema operatiu.
- Atesa la seva popularitat cada cop més gran, els dissenyadors de CPU hi afegeixen cada cop més característiques per facilitar-ne el suport.
- Aquesta arquitectura de virtualització té associada una complexitat alta. Per a determinades aplicacions és més útil la virtualització a nivell del sistema operatiu. Vegem-ho!

Virtualització a nivell d'SO

En un sistema operatiu habitual, una aplicació pot disposar de tots els recursos que ofereix el sistema. Aquests inclouen

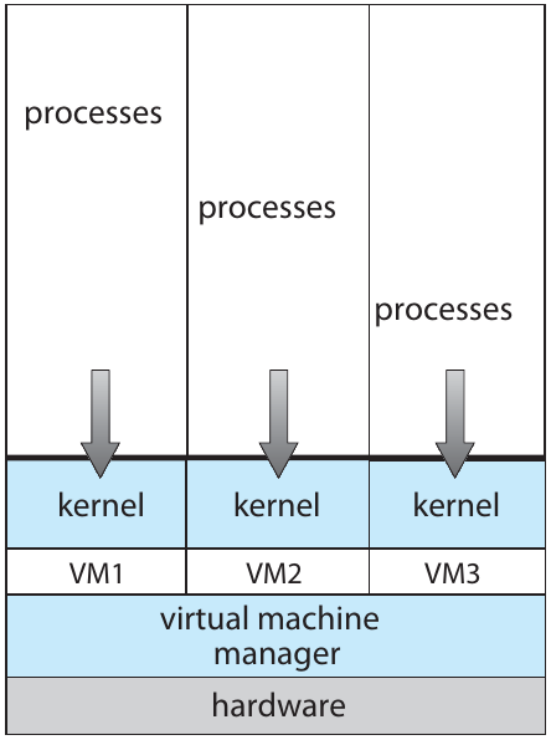
- Utilització de les CPUs i la connexió de xarxa
- Llegir o escriure fitxers, estiguin en disc local o en disc de xarxa
- Altres dispositius com la impressora, la webcam, un fax, ...

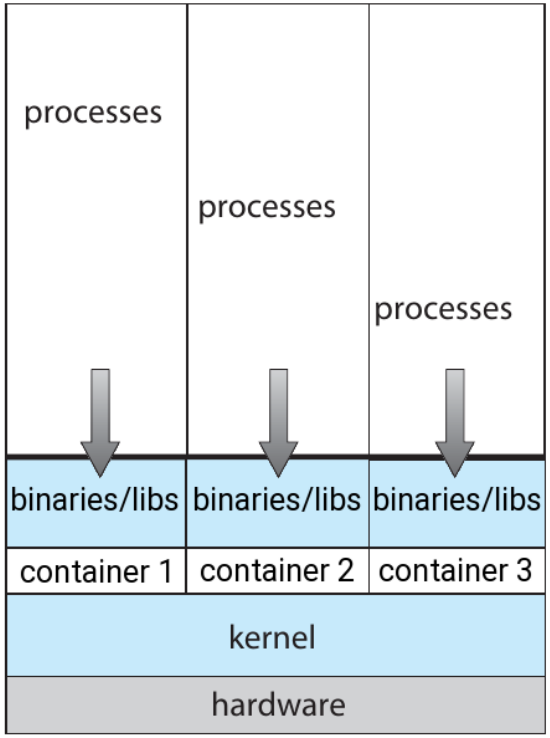
El sistema operatiu permet que un usuari pugui accedir o no a determinats recursos de la màquina. Típicament l'administrador del sistema operatiu (el "root") és qui pot configurar la interacció que l'usuari pot tenir amb els recursos.

L'administrador i/o usuari requereixen, cada cop més sovint, tenir un control fí sobre els recursos als quals poden accedir cadascuna de les aplicacions que s'executen a l'ordinador, encara que l'executi el mateix usuari.

Virtualització a nivell d'SO

La virtualització a nivell d'SO permet executar aplicacions en contenidors on es fan visibles els recursos que volem.





Comparem amb les màquines virtuals (VM)

- A cada VM es pot executar un sistema operatiu diferent.
- A la Virtualització a nivell d'SO només hi ha un únic sistema operatiu en el qual hi ha múltiples contenidors.

Virtualització a nivell d'SO

Vegem-ho amb més detall

- Les màquines virtuals permeten virtualitzar el maquinari. L'hypervisor permet que diverses VMs s'executin en una sola màquina. Cada màquina virtual típicament inclou un sistema operatiu, les aplicacions i les llibreries necessàries.
- Els contenidors són una abstracció a la capa d'aplicacions que empaqueta el codi d'una aplicació i les seves dependències. Es poden executar diversos contenidors a la mateixa màquina i es comparteix el nucli del sistema operatiu. Els processos que s'executen dins del contenidor només "veuen" els recursos que s'hi han assignat.
- Els contenidors i les màquines virtuals tenen avantatges d'aïllament i assignació de recursos similars, però funcionen de manera diferent perquè els contenidors virtualitzen el sistema operatiu en lloc del maquinari.

Virtualització a nivell d'SO

Hi pot haver múltiples contenidors

- Els processos que s'executen a l'interior del contenidor només pot "veure" un subconjunt dels recursos disponibles de l'ordinador.
- A l'interior de cada contenidor podem fer disponibles diferents llibreries. Aquestes poden ser inclús incompatibles entre sí.

Avantatges

- Els processos fan crides a sistema i no és necessari el sistema de virtualització com ho és en una màquina virtual.
- Executar aplicacions en contenidors diferents per assegurar seguretat.
- És utilitzat a sistemes d'allotjament: a cada usuari se li assigna un contenidor que té assignats recursos limitats.
- Sistemes com Docker permeten la distribució del contenidor sencer als seus clients, facilitant la seva instal·lació als ordinadors.

Virtualització a nivell d'SO

De forma genèrica, la Virtualització a nivell d'SO (OS-level virtualization) és un paradigma en què el sistema operatiu permet crear espais d'usuaris aïllats per als processos que s'hi executen.

- El sistema més antic és el chroot jail, del 1979...
- A Linux aquestes instàncies s'anomenen contenidors (sistema LXD, Docker)
- Altres sistemes equivalents com els virtual kernels (DragonFly BSD), Zones (Solaris), ...

Virtualització a nivell d'SO

Els contenidors han evolucionat incloent més característiques de forma que, en cada contenidor, es pot controlar l'espai d'usuari en què s'executen els processos. Entre altres coses, es poden controlar

- El nombre màxim de CPUs que poden utilitzar
- La memòria RAM màxima que poden utilitzar
- El nombre màxim de processos que es poden executar a un contenidor
- L'espai de directoris/fitxers als quals poden accedir
- Moltes altres coses...

Virtualització a nivell d'SO

La tecnologia que permet crear contenidors es basa en la funcionalitat proveïda pels cgroups i namespaces del sistema operatiu.

1. Control groups (cgroups): és la tecnologia del sistema operatiu que permet limitar els recursos d'un conjunt de processos.
2. Namespaces: és la tecnologia que permet aïllar els contenidors entre sí (dins de cada contenidor els processos tindran els seus propis PIDs, cada contenidor tindrà el seu propi cgroup, ...)

Aquesta tecnologia ha estat adoptada per desenvolupar eines que permetin manipular contenidors com, per exemple

- LxC i LxD: De "baix nivell", pels administradors de sistema. Especialitzat per desplegar màquines virtuals (linux) totalment funcionals – `https://linuxcontainers.org/`
- Docker: Orientat al desenvolupador. Especialitzat per desplegar aplicacions: cada contenidor està dissenyat per contenir una aplicació – `https://www.docker.com/`

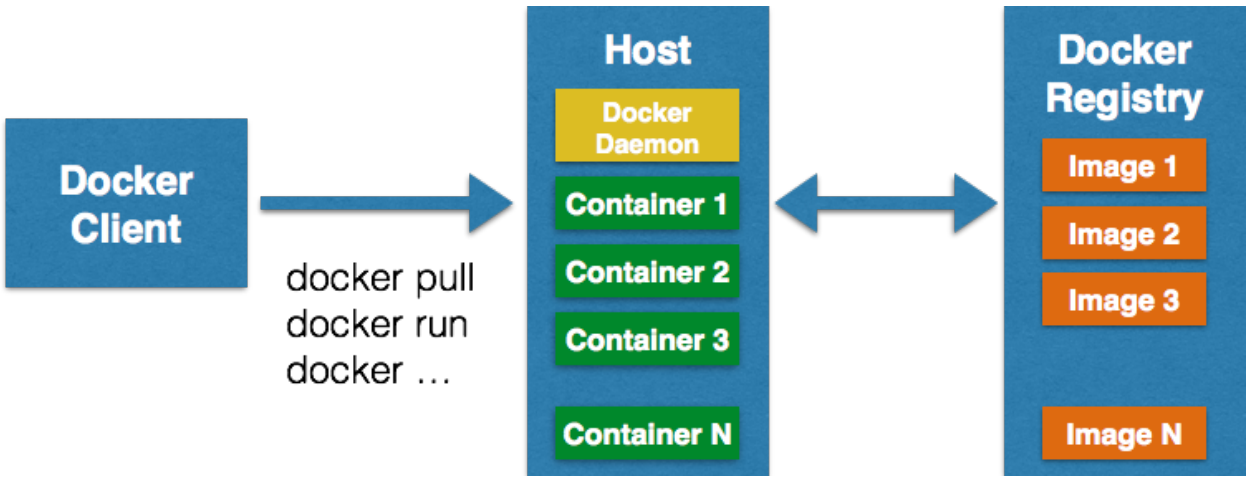
Docker

Internament, Docker utilitza la tecnologia dels cgroups i namespaces per proveir els contenidors. Docker afegeix múltiples característiques per manipular i executar contenidors, de forma equivalent a les eines que es disposen a l'hora de programar:

1. Docker assegura la portabilitat dels seus contenidors entre màquines que utilitzen Docker. Això permet, per exemple, que es distribueixi en un contenidor una aplicació amb totes les llibreries que li fan falta per executar.
2. Docker permet crear contenidors de forma senzilla i inclou funcionalitats de gestió de versions similars a git. Veiem-ho!

Docker

- El Docker registry és un repositori on es pengen els fitxers que permeten generar els contenidor, les Docker images. Una imatge pot ser un fitxer que inclou un compilador Java.
- Un contenidor Docker és una imatge Docker en execució. Per exemple, en executar la imatge amb el compilador Java podrem compilar codi per aquella versió del compilador.
- Docker utilitza una arquitectura client-servidor (Client-Daemon), cosa que permet que el client i el servidor no s'executin a la mateixa màquina.



Alternatives al Docker

Docker és conegut per la seva facilitat d'ús. No és, però, la única eina que permet manipular contenidors

- Singularity, vegeu `https://sylabs.io/docs/`
- Distribució de Linux orientada a executar contenidors rkt, vegeu `https://coreos.com/` que utilitza la tecnologia de contenidors rkt `https://coreos.com/rkt/docs/latest/rkt-vs-other-projects.html`[4].

[4] El contenidor rkt és un tipus de contenidor que s'administra amb una eina diferent del Docker.

Empreses que ofereixen serveis de contenidors

Hi ha moltes empreses que ofereixen serveis per executar contenidors (llistat any 2020). De forma equivalent a les màquines virtuals, utilitzar un servei extern permet escalar fàcilment:

- Google Container Engine
- Amazon EC2 Container Service
- Azure Container Service
- Acquia
- Digital Ocean

Conclusions

La virtualització permet emular un ordinador

- Hem vist dues formes d'aconseguir-ho, a) les màquines virtuals, b) la virtualització a nivell de sistema operatiu.
- Es pot executar un sistema operatiu diferent a cada màquina virtual. Les màquines virtuals permeten limitar els recursos assignats al sistema operatiu.
- Els contenidors s'executen pel sistema operatiu host. A cada contenidor podem limitar els recursos que hi assignem. Les crides a sistema es fan directament al sistema operatiu host.