

Operacijski sistemi



Virtualizacija

Vsebina

- Virtualizacija
- Lastnosti virtualizacije
- Sistemska virtualizacija

Virtualizacija

- Virtualizacija
 - mehanizem, ki nekaj ustvari navidezno
 - **navidezna naprava:** procesor, pomnilnik
 - **navidezni stroj** (celoten sistem, strojna oprema)
 - preslikava navideznega v realni sistem
 - vmesnik in viri navidezne naprave se preslikajo na vmesnik in vire realne naprave, ki implementira navidezno

Virtualizacija

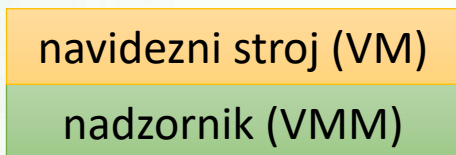
- Primer: navidezni disk oz. datotečni sistem
 - uporabniku enovito dostopen preko klasičnih sistemskih klicev (open, read, write, close)
 - v ozadju pa se navidezni disk preslika na dejanski sistem, npr.:
 - več diskov
 - porazdeljeni datotečni sistem
 - mrežni datotečni sistem
 - jedrni datotečni sistem (npr. /proc)

Virtualizacija

- Abstrakcija in virtualizacija
 - podobna koncepta
 - oba nudita nek vmesnik do navideznega sistema
 - abstrakcija skrije podrobnosti
 - npr. abstrakcija datoteke skrije podrobnosti, kje in kako je datoteka dejansko shranjena
 - virtualizacija lahko podrobnosti pusti odprte
 - npr. navidezni procesor je lahko ravno tako zahteven za programiranje kot dejanski

Virtualizacija

- Virtualizacijska programska oprema
 - programska oprema za virtualizacijo
 - vključuje tudi upravljanje in konfiguracijo
 - nudi virtualizacijski vmesnik navedeznim strojem



gost (guest)

- sistem, ki ga virtualiziramo
- navidezni stroj

gostitelj (host)

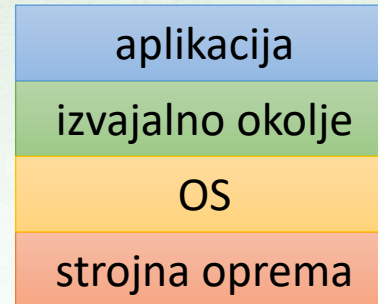
- dejanski sistem, ki poganja virtualizacijsko programsko opremo



Virtualizacija

- Vrste virtualizacije

- procesni navidezni stroj
 - virtualizacijska PO: izvajalno okolje (**runtime**)
 - uporablja domači (native) OS in strojno opremo
 - omogoča prenosljivost aplikacij
- sistemski navidezni stroj
- vsebniki



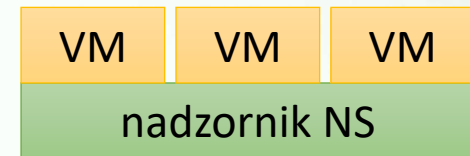
JVM, CLR,
Parrot, Neko,
Lua, Python



Lastnosti virtualizacije

- Replikacija virov

- videz da je več virov (naprav), kot jih dejansko je
 - omogoča sočasnost navideznih strojev (in OS in aplikacij)



- Varnost in neodvisnost

- navidezni stroji so med seboj neodvisni in izolirani
- zaščita gostiteljskega OS pred aplikacijami
 - dva nivoja zaščite
- zaščita nadzornika pred gostitelji
 - dodaten virtualizacijski nivo zaščite

Lastnosti virtualizacije

- Prilagodljivost in enostavno upravljanje
 - večja razpoložljivost in zanesljivost
 - selitev strojev med virtualizacijskimi platformami
 - zmanjšanje stroškov
 - več navideznih strojev na enen dejanskem stroju
 - lažje upravljanje (navideznih) strojev
 - namestitev OS, razhroščevanje
 - podpora za starim OS

Lastnosti virtualizacije

- Slabosti

- manjša učinkovitost
 - na dejanskem stroju teče več navideznih
- upravljanje dejanskega stroja je bolj zahtevno
 - namestitev in konfiguracija nadzornika
- virtualizacija lahko ni popolna
 - dostopnost dejanskih naprav v navideznem stroju
 - niso dostopne, delno dostopno, drugačen vmesnik

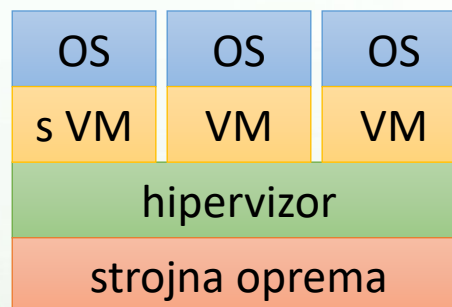
Sistemska virtualizacija

- Sistemski navidezni stroj
 - virtualizira celoten sistem oz. strojno opremo
 - procesor, pomnilnik, naprave itd.
 - stara ideja – zakaj se ni razširila že v 60tih?
 - veliki osrednji računalniki niso bili razširjeni
 - mali računalniki so bili poceni
 - nova ideja – zakaj se je razširila šele zdaj?
 - podatkovni centri so postali ogromni
 - nudi boljšo izkoriščenost strežnikov
 - zahtevna administracija velike količine strežnikov
 - visoki stroški obratovanja strežnikov
 - omogoča varno souporabo strežnikov

IBM

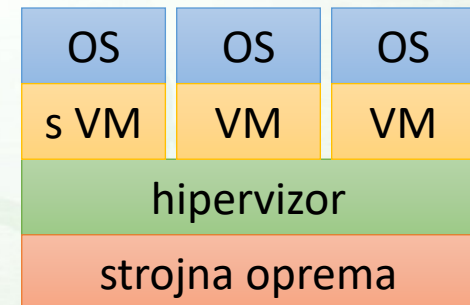
Sistemska virtualizacija

- Virtualizacija tipa 1 (bare metal, hypervisor-based)
 - nadzornik navideznih naprav (VM monitor)
 - hipervizor (hypervisor)
 - teče **neposredno** na strojni opremi
 - servisni navidezni stroj
 - omogoča konfiguracijo in upravljanje sistema
 - privilegirane operacije, dostop do strojne opreme
 - poganja gonilnike za strojno opremo



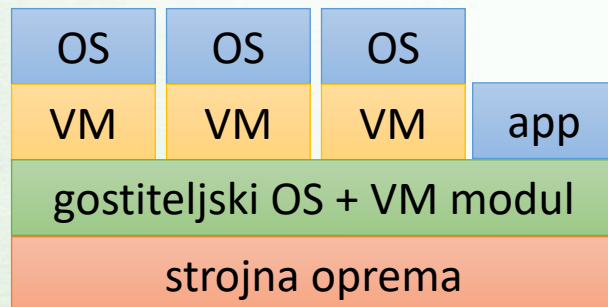
Sistemska virtualizacija

- Virtualizacija tipa 1
 - Xen (odprtokodni), Citrix XenServer
 - domene dom0, domUs
 - gonilniki tečejo v dom0
 - VMware ESXi
 - gonilniki so del nadzornika
 - proizvajalci morajo dobaviti gonilnike
 - odprt API, remote API management
 - Microsoft Hyper-V



Sistemska virtualizacija

- Virtualizacija tipa 2
 - temelji na gostiteljskem OS (host-OS-based)
 - primeri: VMware Workstation/Player/Fusion, Oracle VirtualBox, Microsoft Virtual PC, ...



Sistemska virtualizacija

- Polna virtualizacija (full virtualization)
 - preslikava vseh pomembnih funkcij dejanskega stroja v funkcije navideznih strojev
 - celoten nabor strojnih ukazov, prekinitve, vhodno/izhodne operacije, dostop do pomnilnika
 - gostujoči OS se ne zaveda virtualizacije in deluje enako kot v ne virtualiziranem okolju

Sistemska virtualizacija

- Paravirtualizacija

- učinkovitost za ceno spremembe gostujočega OS
- gostujoči OS ve, da se izvaja virtualizirano
- OS ne izvaja privilegiranih ukazov
- ampak eksplicitno kliče nadzornika
 - nadzorni klici, hiper klici (hyper calls)
- originalno
 - Xen, sedaj XenSource

Izvedba virtualizacije

- Osnovna pristopa
 - souporaba/deljenje časa (time sharing/partitioning)
 - gre za deljenje procesorskega časa
 - souporaba/deljenje prostora (space sharing/partitioning)
 - gre za deljenje fizičnega pomnilnika

Izvedba virtualizacije

- Podpora strojne opreme
 - Intel VT-x in AMD-V
 - root in non-root način delovanja
 - obroč 3: non-root nivo 3, aplikacije
 - obroč 0: non-root nivo 0, gostujoči operacijski sistem
 - obroč -1: root nivo 0, nadzornik navideznih strojev
 - preklop med gostujočim OS in nadozrnikom
 - VM-exit - past
 - ko gostujoči OS zahteva privilegirano operacijo
 - nadzor prevzame nadzornik navideznih strojev
 - VM-entry
 - izhod iz nadzornik oz. vstop nazaj v navidezni stroj

Izvedba: virtualizacija procesorja

- Virtualizacija procesorja
 - strojni ukazi se izvedejo neposredno na procesorju
 - ne-privilegirani ukazi: učinkovitost izvajanja
 - privilegirani ukazi: past, ki jo rokuje nadzornik
 - izvedba privilegiranih ukazov
 - emulira jih nadzornik, tako kot to pričakuje gostujoči OS

Izvedba: virtualizacija procesorja

- Težave pri izvedbi virtualizacije
 - primer Intel x86, pred 2005
 - brez root načina
 - obroč: 0: nadzornik, 1: OS, 3: aplikacije
 - 17 privilegiranih ukazov ni prožilo pasti v obroč 1+
 - STI, CLI, POPF, PUSHF, ...
 - le potihoma niso uspeli
 - nadzornik torej ni naredil nič
 - OS je mislil, da je vse ok

Izvedba: virtualizacija procesorja

- Prevajanje (binary translation)
 - v kodi nadomesti problematične ukaze
 - cilj: brez sprememb v gostujočem OS
 - dinamično dvojiško prevajanje
 - pregledovanje blokov kode, ki se bodo izvedli
 - nadomeščanje privilegiranih ukazov
- pionirji
 - Mendel Rosenblum et al., Stanford -> Vmware
 - nagrada za “reinventing virtualization”

Izvedba: virtualizacija pomnilnika

- Polna virtualizacija



- 1. možnost: dvojno preslikovanje
 - programsko + strojno (MMU, TLB)
 - neučinkovito
- 2. možnost: senčna preslikovalna tabela
 - na podlagi preslikovalnih tabel gostujočih OS
 - nadzornik vodi preslikovalno tabelo, ki slika iz navideznih naslovov v dejanske fizične naslove (MMU)

Izvedba: virtualizacija pomnilnika

- Paravirtualizacija

- s preslikovalnimi tabelami upravlja nadzornik
- gostujoči OS se zaveda virtualizacije
 - lahko prilagodi uporabo pomnilnika
 - pri nadzorniku eksplicitno registrira svoje preslikovalne table

Izvedba: virtualizacija naprav

- Izzivi
 - veliko različnih naprav
 - manj standardizacije obnašanja naprav



primerjava s CPU
in pomnilnikom

Izvedba: virtualizacija naprav

- Pristopi

- passthrough model

- VMM konfigurira dostop do naprav
 - navidezni stroj pridobi ekskluzivni in neposredni (zaobide nadzornika) dostop do naprave
 - souporaba naprave zelo težavna
 - migracija navideznih strojev je težavna

Izvedba: virtualizacija naprav

- Hypervisor-direct model
 - nadzornik prestreže vse dostope do naprave
 - emulira odziv naprave
 - generična V/I operacija
 - uporabi gonilnik (znotraj nadzornika)
 - nadzorni stroj je ločen od naprave
 - migracija in souporaba je lažja
 - VMware ESXi
 - zakasnitev operacij dostopa do naprave
 - hipervizor mora omogočati gonilnike

Izvedba: virtualizacija naprav

- Split-device driver model
 - dostop do naprave je razdeljen med
 - front-end gonilnik v navideznem stroju
 - back-end gonilnik v servisnem navideznem stroju (tip 1) ali v gostiteljskem OS (tip 2)
 - omejeno na paravirtualizirane goste
 - ni režije zaradi emulacije