SO Curs 10

Virtualizarea este asigurata de un Virtual Machine Monitor poate rula:

- direct pe hardware
- ca o aplicatie in sistemul gazda

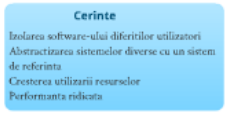
Functii VMM

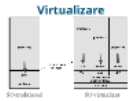
virtualizare resurse (CPU, memorie, disk, retea)
gestionare evenimente
alocare resurse

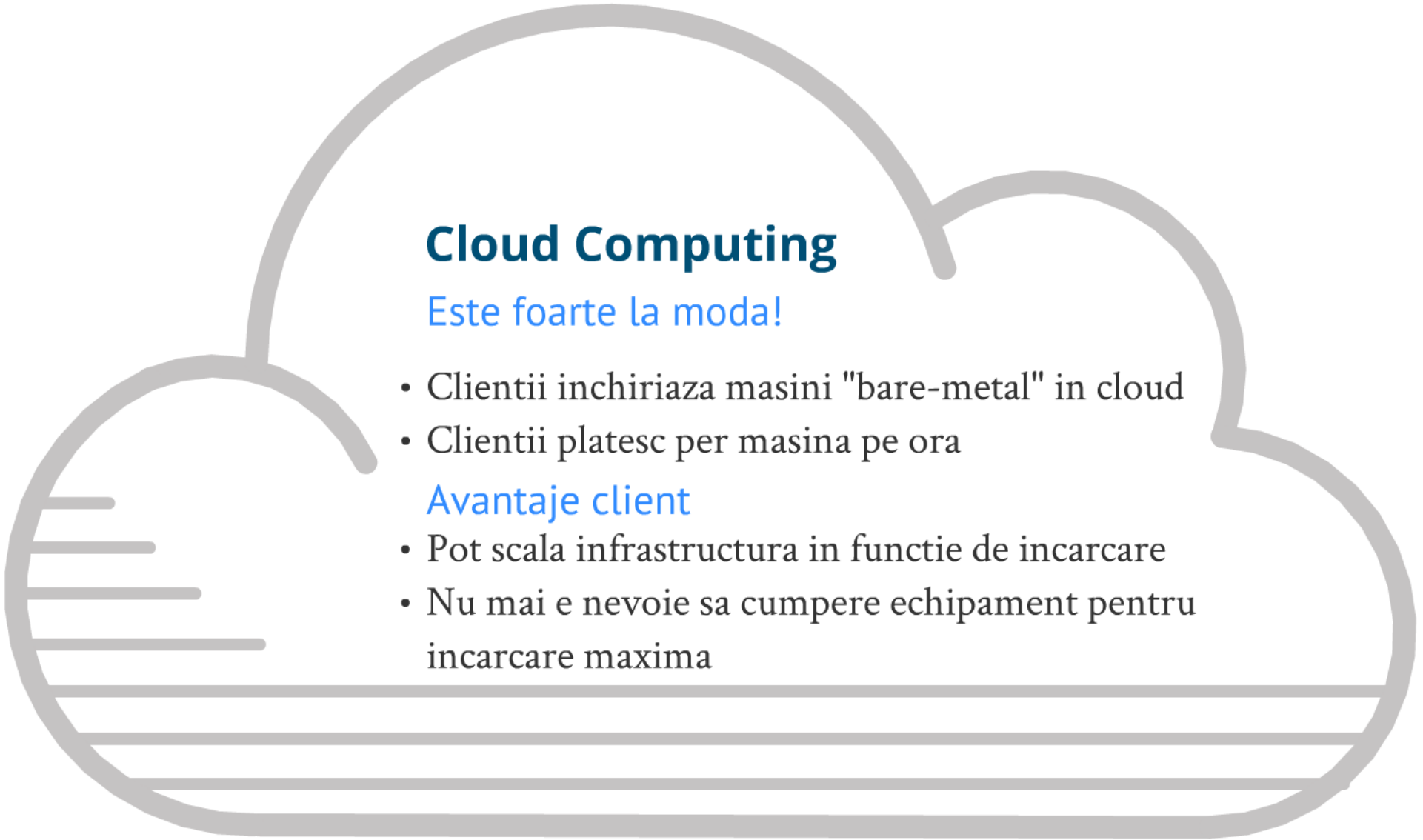
Categorii de VMM

Full Virtualization
ParaVirtualization
OS-level virtualization

Virtualizare







Cloud Computing

Este foarte la moda!

- Clientii inchiriaza masini "bare-metal" in cloud
- Clientii platesc per masina pe ora

Avantaje client

- Pot scala infrastructura in functie de incarcare
- Nu mai e nevoie sa cumpere echipament pentru incarcare maxima

Haideti sa ne facem un cloud!

Idee: cumparam multe masini si dam fiecarui client acces la masini la cerere

Dar costa cam mult!

Observatie

Un server al unui client este incarcat in medie 30%

Clientii nu folosesc masinile simultan

Ce ar fi daca....

Am servi mai multi clienti cu aceeasi masina fizica?

- am putea creste utilizarea masinilor fizice
- am reduce costurile

SO Curs 10

Virtualizarea este asigurata de un Virtual Machine Monitor poate rula:

- direct pe hardware
- ca o aplicatie in sistemul gazda

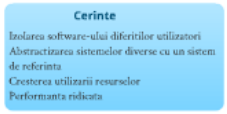
Functii VMM

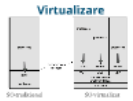
virtualizare resurse (CPU, memorie, disk, retea)
gestionare evenimente
alocare resurse

Categorii de VMM

Full Virtualization
ParaVirtualization
OS-level virtualization

Virtualizare





Cerinte

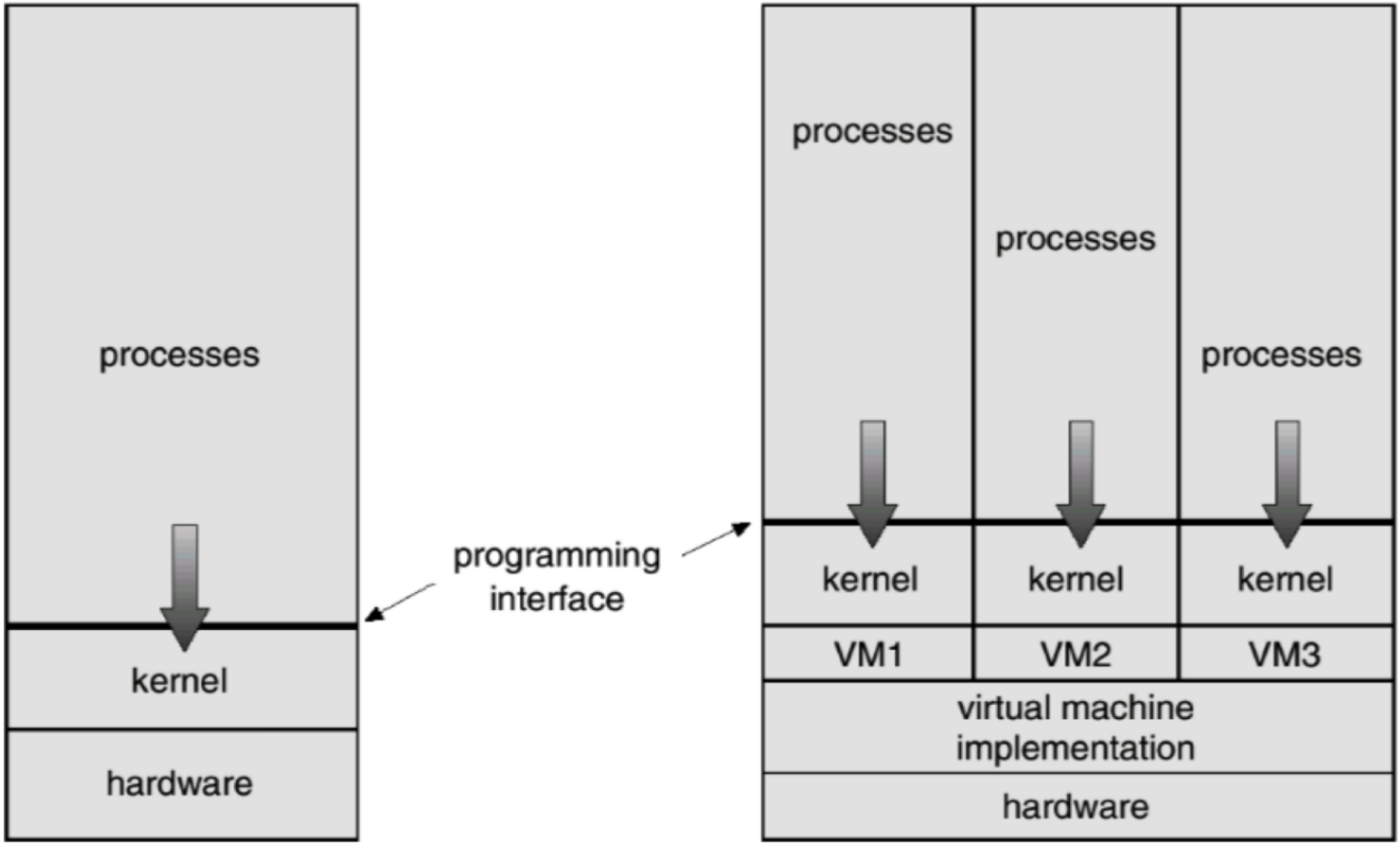
Izolarea software-ului diferitilor utilizatori

Abstractizarea sistemelor diverse cu un sistem de referinta

Cresterea utilizarii resurselor

Performanta ridicata

Virtualizare



SO traditional

SO virtualizat

Cum putem folosi virtualizarea?

Consolidare
mai multe masini virtuale pe aceeasi masina fizica

Sandboxing
rulam cod "buggy" intr-un mediu izolat

Development
putem testa si depana noi SO mai usor

Domenii de executie multiple

Cerinte pentru implementare

Enuntate de Popek si Goldberg in 1974

Echivalenta	Un program care ruleaza intr-un VMM trebuie sa se comporte identic cu o rulare direct pe hardware
Siguranta	Virtual Machine Monitor trebuie sa controleze resursele virtualizate
Eficienta	Un procent semnificativ de instructiuni trebuie sa fie executate fara interventia VMM.

O arhitectura are trei categorii de instructiuni

- Instructiuni privilegiate - transfera controlul VMM
- Instructiunile sensibile - afecteaza functionarea VMM
- Instructiuni neprivilegiate - se executa nativ

Teorema [Popek si Goldberg]

Un sistem poate fi virtualizat daca setul de instructiuni sensibile ale sistemului este un subset al instructiunilor privilegiate

Putem virtualiza x86?

- Instructiunile sensibile - afecteaza functionarea VMM
- Instructiuni neprivilegiate - se executa nativ

Teorema [Popek si Goldberg]

Un sistem poate fi virtualizat daca setul de instructiuni sensibile ale sistemului este un subset al instructiunilor privilegiate

Putem virtualiza x86?

Contine 17 instructiuni sensibile si neprivilegiate
e.g. pushf, popf

Virtualizarea este asigurata de un Virtual Machine Monitor poate rula:

- direct pe hardware 
- ca o aplicatie in sistemul gazda 

Functii VMM

virtualizare resurse (CPU, memorie, disk, retea)
gestionare evenimente
alocare resurse

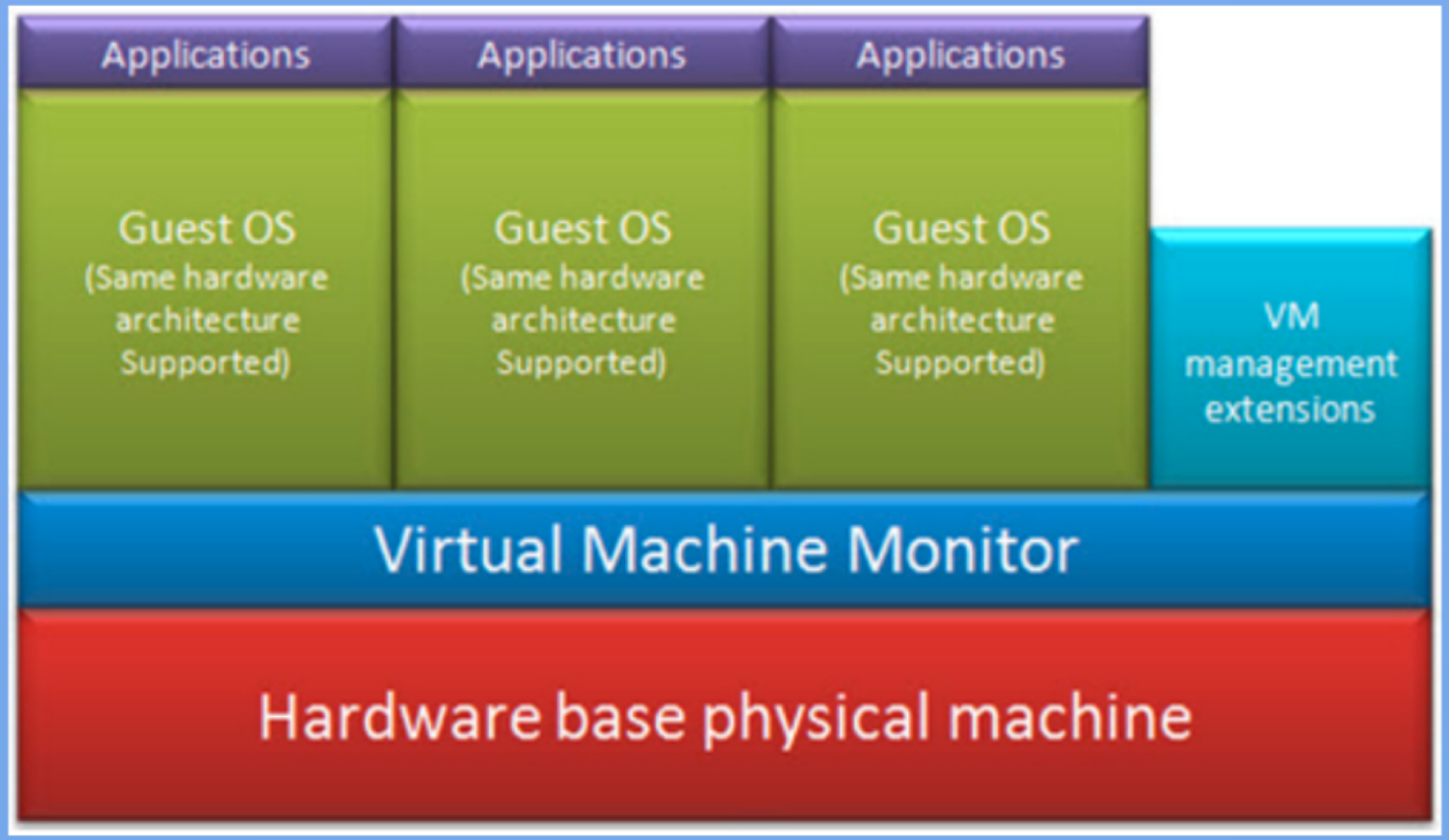
Categorii de VMM

Full Virtualization
Paravirtualization
OS-level virtualization

Virtualizare

Cum putem folosi virtualizarea?

Cerinte pentru implementare


Applications
Applications
Applications
Guest OS
(Same hardware architecture Supported)
Guest OS
(Same hardware architecture Supported)
Guest OS
(Same hardware architecture Supported)
VM management extensions
Virtual Machine Monitor
Hardware base physical machine

Virtualizarea este asigurata de un Virtual Machine Monitor poate rula:

- direct pe hardware 
- ca o aplicatie in sistemul gazda 

Functii VMM

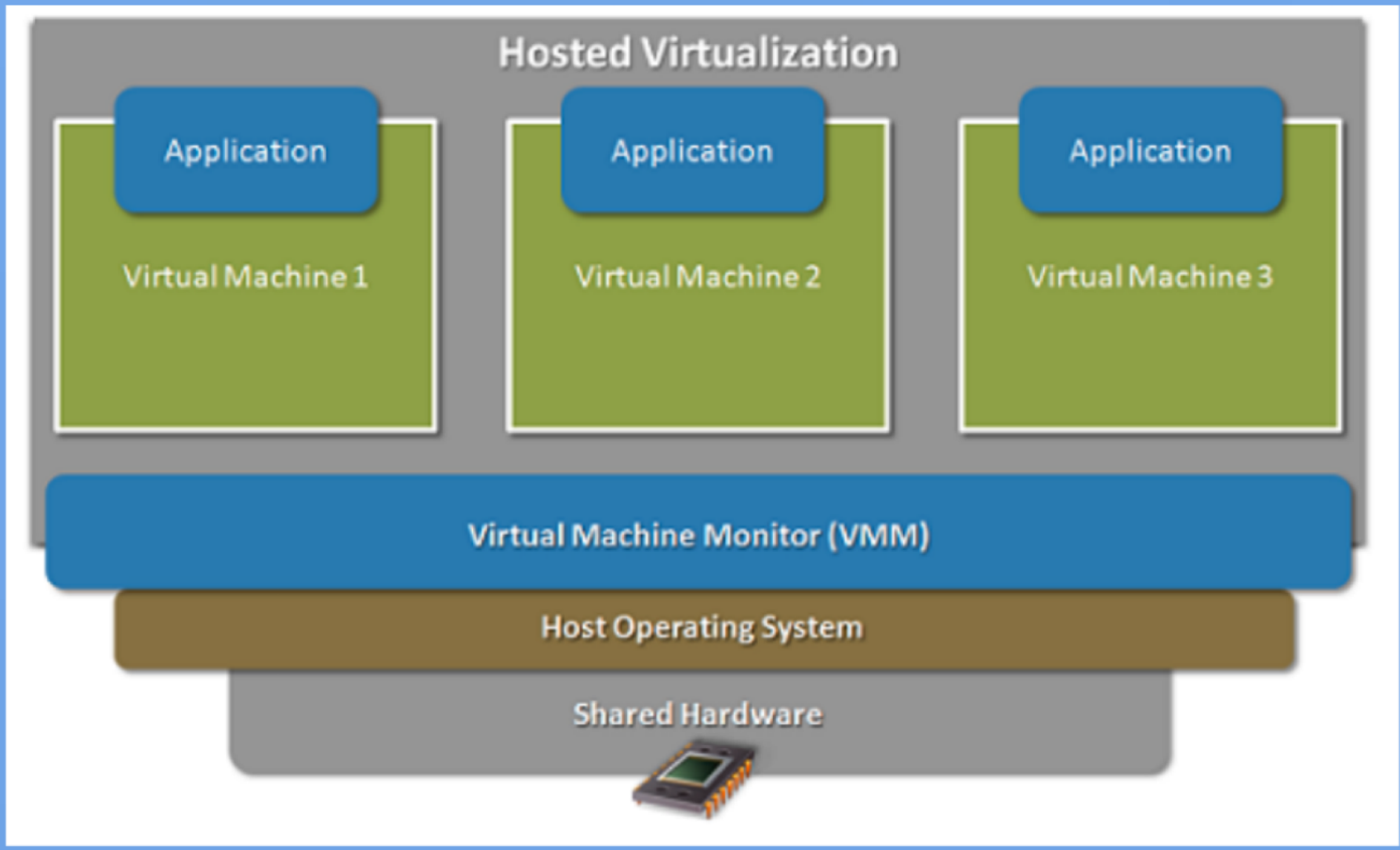
virtualizare resurse (CPU, memorie, disk, retea)
gestionare evenimente
alocare resurse

Categorii de VMM

Full Virtualization

Paravirtualization

OS-level virtualization


Hosted Virtualization
Application
Application
Application
Virtual Machine 1
Virtual Machine 2
Virtual Machine 3
Virtual Machine Monitor (VMM)
Host Operating System
Shared Hardware

Virtualizarea este asigurata de un Virtual Machine Monitor poate rula:

- direct pe hardware 
- ca o aplicatie in sistemul gazda 

Functii VMM

virtualizare resurse (CPU, memorie, disk, retea)
gestionare evenimente
alocare resurse

Categorii de VMM

Full Virtualization
Paravirtualization
OS-level virtualization

Virtualizare

Cum putem folosi virtualizarea?

Cerinte pentru implementare

O arhitectura are trei categorii de instructiuni

Functii VMM

virtualizare resurse (CPU, memorie, disk, retea)

gestionare evenimente

alocare resurse

Categorii de VMM

Full Virtualization

Paravirtualization

OS-level virtualization

Simuleaza complet hardware-ul

Aplicatiile (SO) ruleaza nemodicate peste VMM

Satisfac cerintele Popek & Goldberg

x86 nu indeplineste criteriile Popek & Goldberg fara extensii recente (Intel VT, AMD-V)

Putem virtualiza complet x86?

Probabil ca da! VMWare, Parallels, Virtual Box

VMWare

Virtualizeaza ia32 prin binary translation:
Rescrie instructiunile sensibile si neprivilegiate

Ofera seturi generice de device-uri:
tastatura si moust PS/2, floy, controller IDE, CD-ROM, etc

VMDriver ofera acces mai rapid la device-uri

Virtualizare CPU

Detecteaza basic blocks
Pentru un basic bloc:

- Inlocuieste instructiuni sensibile/privilegiate cu apel VMM
- Rescrie ultima instructiune
- Lanseaza in executie

Mentine cache de basic blocks pentru a creste viteza

Virtualizare retea

Placa de retea este pusa in mod promiscuous
Creeaza un bridge in software care primeste toate pachetele si le ruteaza catre:

- SO gazda sau
- SO guest

Virtualizare CPU

Detecteaza basic blocks

Pentru un basic bloc:

- Inlocuieste instructiuni sensibile/privilegiate cu apel VMM
- Rescrie ultima instructiune
- Lanseaza in executie

Mentine cache de basic blocks pentru a creste viteza

Virtualizare retea

Placa de retea este pusa in mod promiscuous

Creeaza un bridge in software care primeste toate pachetele si le ruteaza catre:

- SO gazda sau
- SO guest

Categorii de VMM

Full Virtualization

Paravirtualization

OS-level virtualization

Virtualizarea completa costa performanta

e.g. Xen, VMWare ESX

Ofera o interfata diferita SO guest

SO trebuie modificate (minim)

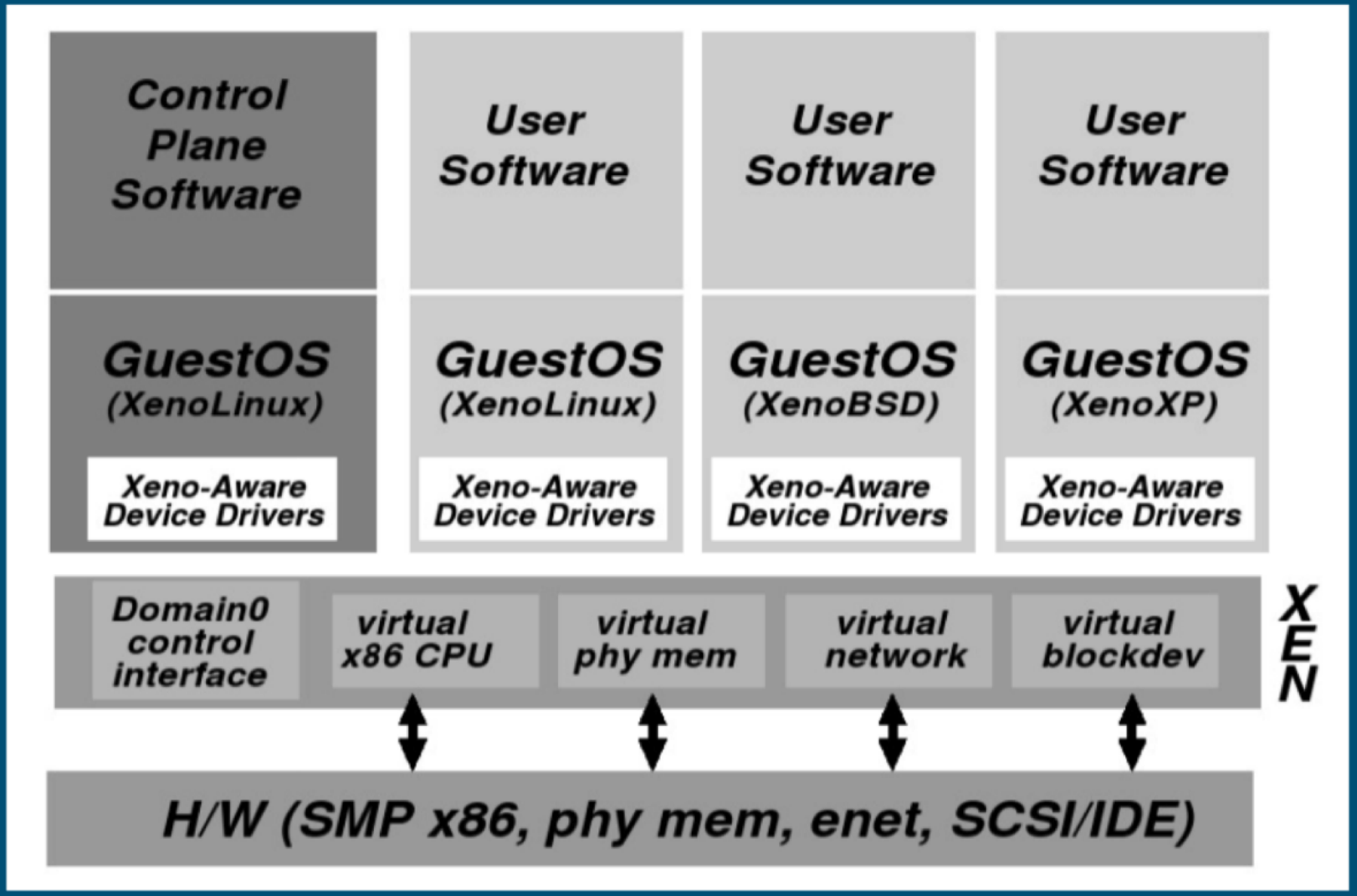
Nu schimbam aplicatiile

Open souce (x86, x86-64, PPC)

Dom0 - Linux sau NetBSD

Guest - Linux, Minix, *BSD, Windows XP, etc

XEN



XEN Details

CPU

x86 has 4 rings:

- Kernels run in ring 0
- Apps run in ring 3

The hypervisor takes over ring 0

- Export a hypercall API (eq. syscall)
- Traps privileged instructions
- The OS moves to ring 1

Exceptions

Page faults, system calls, general protection faults, etc. must go through XEN:

- it has a handler table
- these call the guest's handlers if needed
- fast path for system calls: inherit the guest's handlers directly

Memory management

XEN mentine pentru fiecare frame un reference counter si un tip:

- PD - page directory
- PT - page table
- GDT - global descriptor table
- LDT - local descriptor table
- R/W

Modificarile in paginile de tip non-RW sunt supravegheate
Sunt permise numai prin hypercall, si validate de XEN

XEN ocupa 64MB la inceputul oricarui spatiu de adresa. De ce?

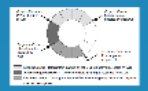
Fiecare guest isi gestioneaza propria memorie: page tables, segment descriptors

Device I/O

Numai dom0 are acces direct la hardware
GuestOS vad simple abstractii in loc de emulare:

- memorie partajata
- buffere de descriptori circulare, asincrone

Intreruperile sunt tratate de XEN
XEN poate genera evenimente in guest OS



Network I/O Virtual Interfaces, fiecare cu 2 ringuri (Tx si Rx)

Disk I/O Virtual Block Devices mapate de dispozitive fizice

CPU

x86 has 4 rings:

- Kernels run in ring 0
- Apps run in ring 3

The hypervisor takes over ring 0

- Export a hypercall API (eq. syscall)
- Traps privileged instructions
- The OS moves to ring 1

Exceptions

Page faults, system calls, general protection faults, etc. must go through XEN:

- it has a handler table
- these call the guest's handlers if needed
- fast path for system calls: inherit the guest's handlers directly

Memory management

XEN mentine pentru fiecare frame un reference counter si un tip:

- PD - page directory
- PT - page table
- GDT - global descriptor table
- LDT - local descriptor table
- R/W

Modificarile in paginile de tip non-RW sunt supravegheate
Sunt permise numai prin hypercall, si validate de XEN

XEN ocupa 64MB la inceputul oricarui spatiu de adresa. De ce?

Fiecare guest isi gestioneaza propria memorie: page tables, segment descriptors

Device I/O

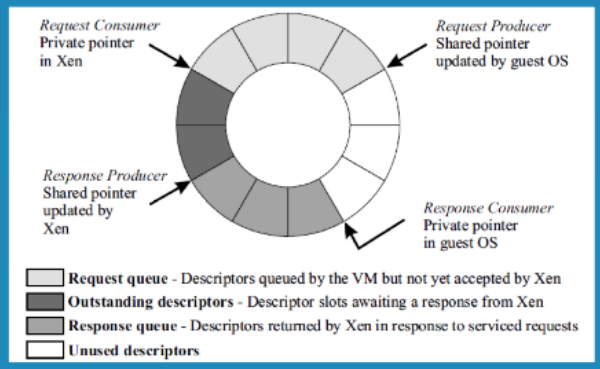
Numai dom0 are acces direct la hardware

GuestOS vad simple abstractii in loc de emulare:

- memorie partajata
- buffere de descriptori circulare, asincrone

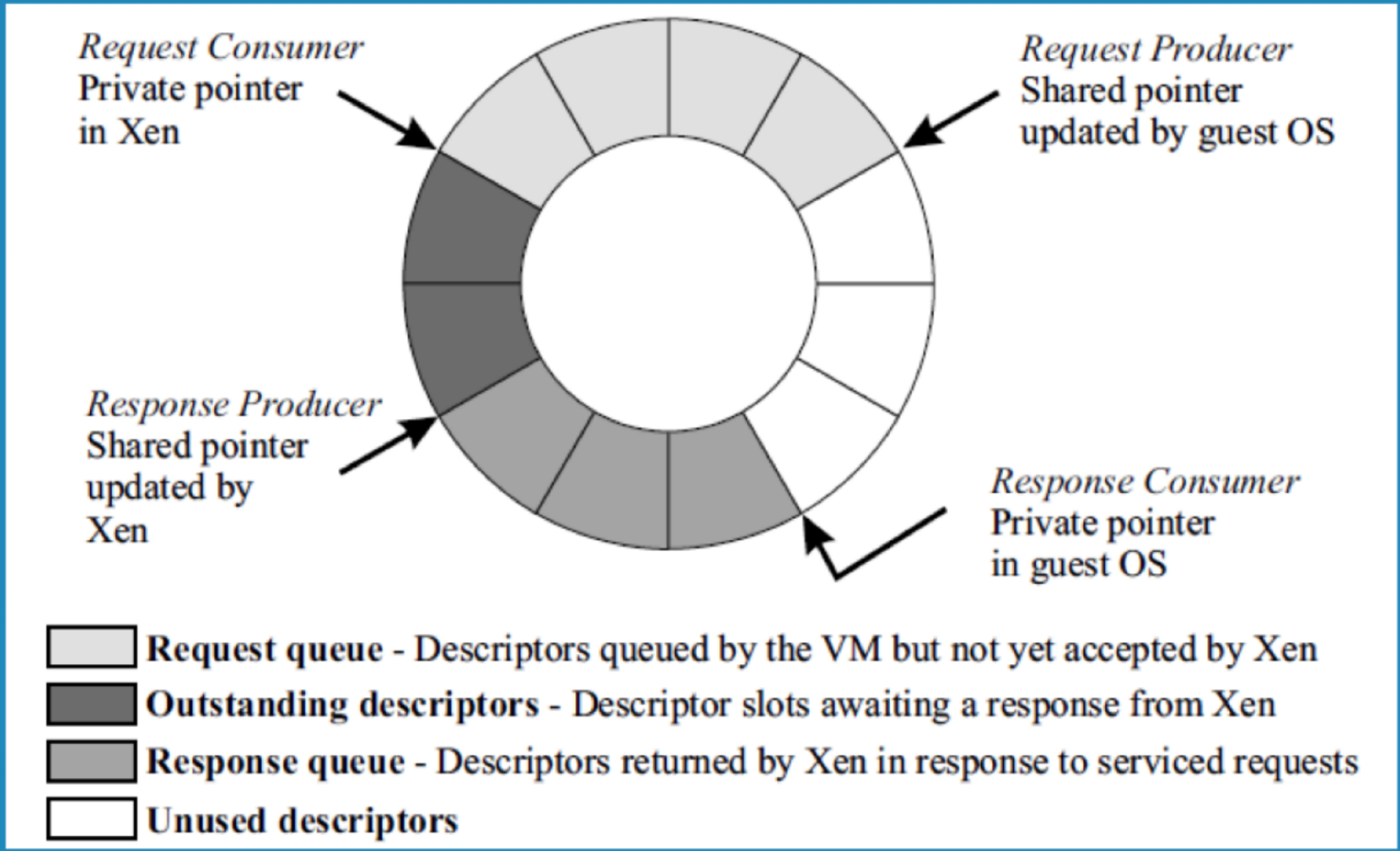
Interuperile sunt tratate de XEN

XEN poate genera evenimente in guest OS



Network I/O Virtual Interfaces, fiecare cu 2 ringuri (Tx si Rx)

Disk I/O Virtual Block Devices mapate de dispozitive fizice


Request Consumer
Private pointer
in Xen
Request Producer
Shared pointer
updated by guest OS
Response Producer
Shared pointer
updated by
Xen
Response Consumer
Private pointer
in guest OS
Request queue - Descriptors queued by the VM but not yet accepted by Xen
Outstanding descriptors - Descriptor slots awaiting a response from Xen
Response queue - Descriptors returned by Xen in response to serviced requests
Unused descriptors

Categorii de VMM

Full Virtualization

Paravirtualization

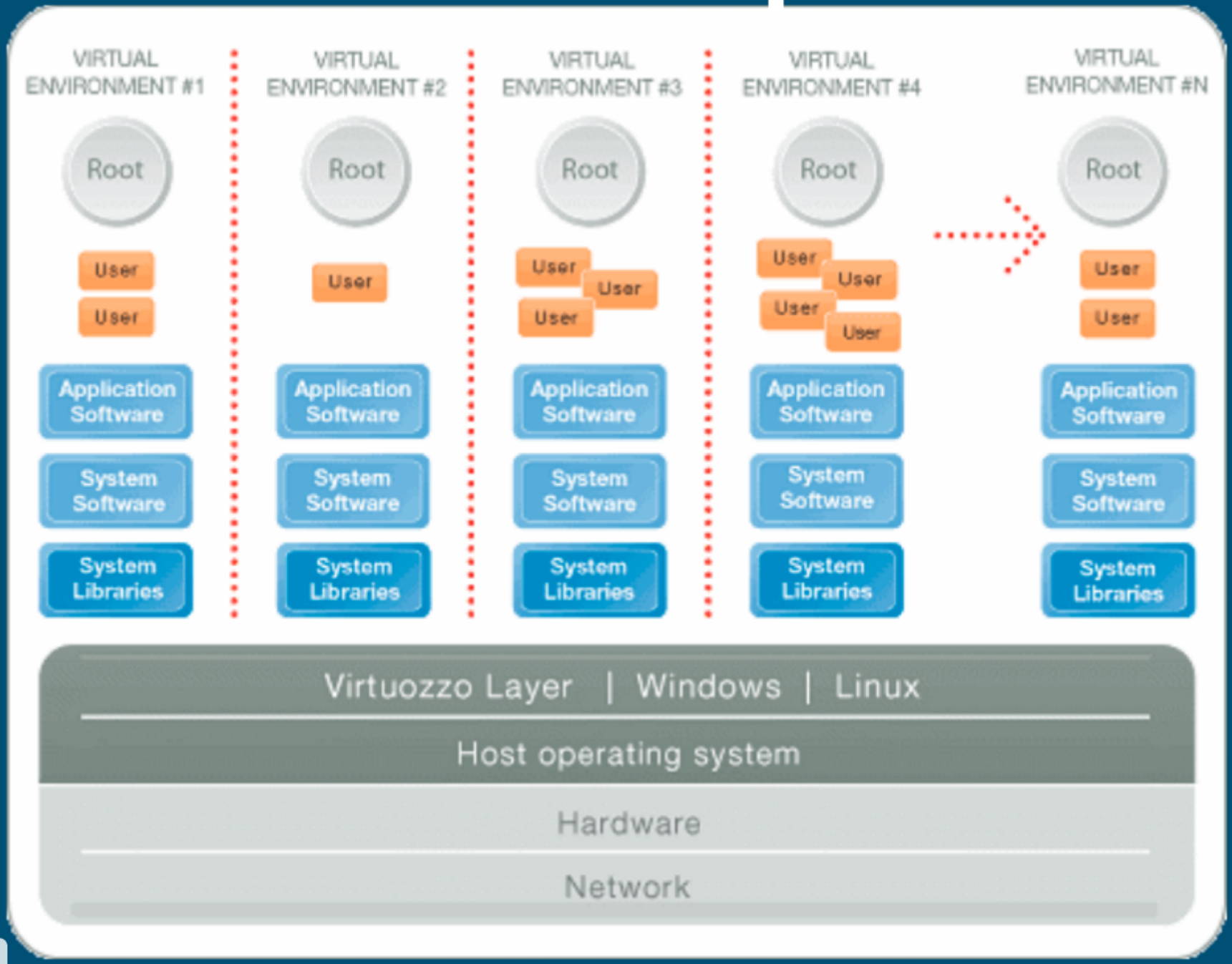
OS-level virtualization

Virtualizare la nivel de SO

Un singur kernel
Mai multe instante de spatii utilizator
Implementare avansata de chroot (FS, procese, I/O, retea, cote, memorie, CPU)

Avantaje	overhead mic
Exemple	OpenVZ, FreeVPS, FreeBSD jail

Virtuozzo/OpenVZ



OpenVZ

Sisteme guest: diferite distributii Linux cu acelasi nucleu

Virtual Environment (VE)

- VPS, container, partition
- mediu izolat de executie
- copie a unui sistem OS Linux (FS, users, retea, tabele rutare, firewall, etc)

Nucleul este un Linux modificat cu:

- virtualizare si izolare
- managementul resurselor
- checkpointing

SO Curs 10

Virtualizarea este asigurata de un Virtual Machine Monitor poate rula:

- direct pe hardware
- ca o aplicatie in sistemul gazda

Functii VMM

virtualizare resurse (CPU, memorie, disk, retea)
gestionare evenimente
alocare resurse

Categorii de VMM

Full Virtualization
Paravirtualization
OS-level virtualization

Virtualizare