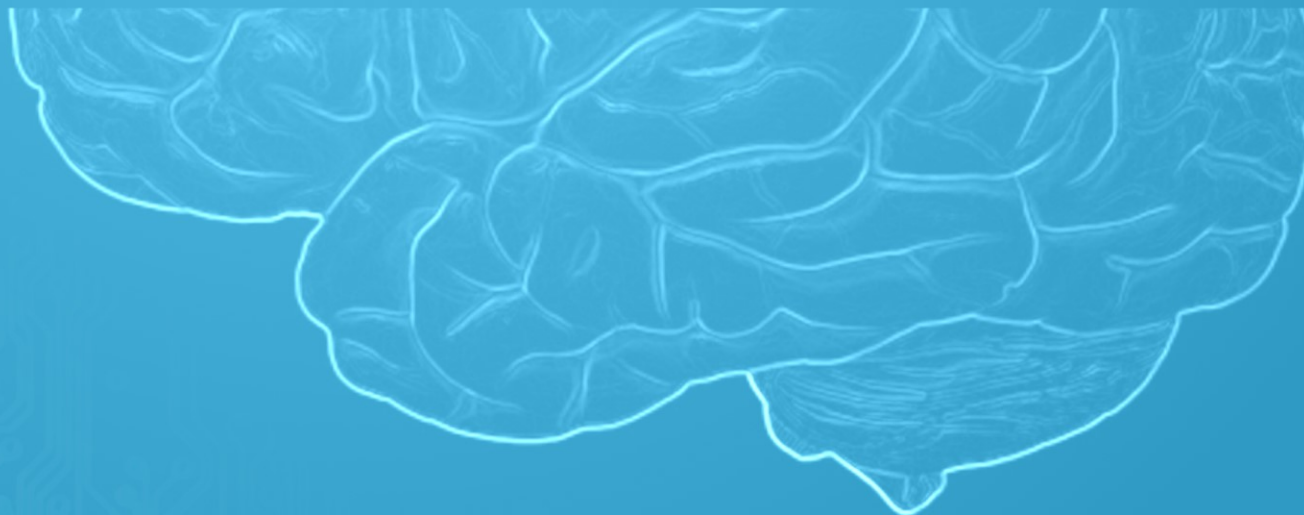




Sisteme Tolerante la Defecte Introducere Cloud Computing

Lect. Dr. Ing. Cristian Chilipirea – cristian.chilipirea@mta.ro







Articol științific principal despre Cloud

Above the Clouds: A Berkeley View of Cloud Computing

Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz,
Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia
(Comments should be addressed to abovetheclouds@cs.berkeley.edu)

UC Berkeley Reliable Adaptive Distributed Systems Laboratory *
<http://radlab.cs.berkeley.edu/>

February 10, 2009

KEYWORDS: Cloud Computing, Utility Computing, I

1

Cloud
IT in
pure
purch
in ha
prov
prov
with
hour
large

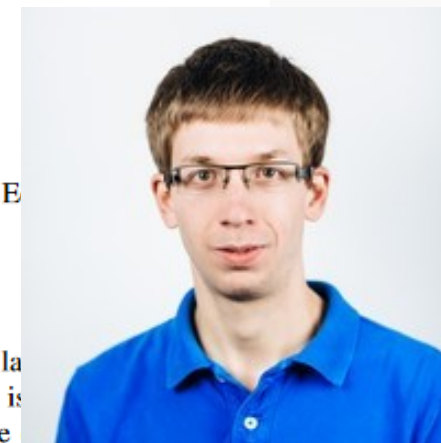


of computing as a
attractive as a ser
ideas for new Inter
ne human expense
ity does not meet t
popular, thus missin
sults as quickly as t
for 1000 hours. Th
y of IT.



System E

transform a la
hardware is
the large
e concerned about over-
ostly resources, or under-
ue. Moreover, companies
sing 1000 servers for one
out paying a premium for





Cloud-ul este despre Cost \$\$\$\$

Economie de scală (cumpără angro, în bulk).

- Cumpăra hardware la un preț mai mic decât ar face asta chiar unele companii mari.

Pay-as-you-go (efectiv resursele se închiriază).

- Noutatea cloud-ului e că acest lucru se întâmplă automat.
- Majoritatea se plătesc la nivel de ore

People are expensive

- Un cluster necesită o echipă 24/7
- Echipa trebuie să aibă oameni cu specializări diverse (hardware, [security](#), networking, etc.)

Ideal pentru **start-up-uri**, echipe **mici**, proiecte **scurte**

- Cost o lună resurse puține cloud << cost cluster



Când NU cloud?



Când NU cloud?

Date **sensibile**

- Poate fi chiar ilegal să pui anumite date (medicale) pe Cloud public

Proiect foarte **mare** cu **durată mare**.

- $\text{cost cluster} < \text{cost multe resurse cloud} * \text{periodă timp}$

Nevoi de bandwidth/latență mari

- Cloud-ul public este peste internet.

Control

- Serviciile și garanțiile se pot schimba în timp



Elasticitate

Elasticity in Cloud Computing: What It Is, and What It Is Not

Nikolas Roman Herbst, Samuel Kounev, Ralf Reussner
Institute for Program Structures and Data Organisation
Karlsruhe Institute of Technology
Karlsruhe, Germany
{herbst, kounev, reussner}@kit.edu

Abstract

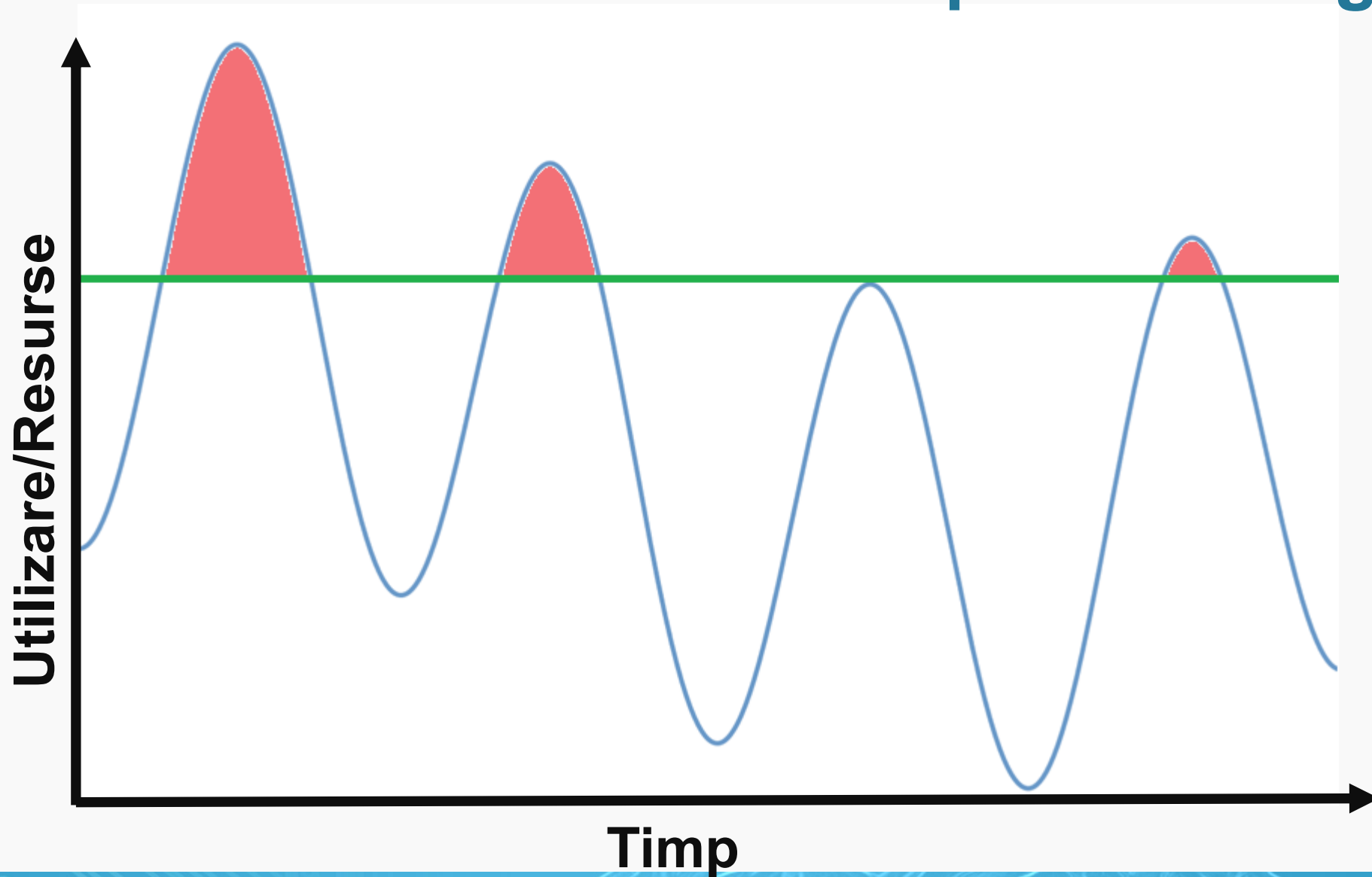
Originating from
term elasticity
of cloud comput

even in the naming of specific products or services. Even
though some definitions are intended to enable standard

Elasticity is the degree to which a system is able to adapt to workload changes by provisioning and de-provisioning resources in an autonomic manner, such that at each point in time the available resources *match* the current demand as closely as possible.

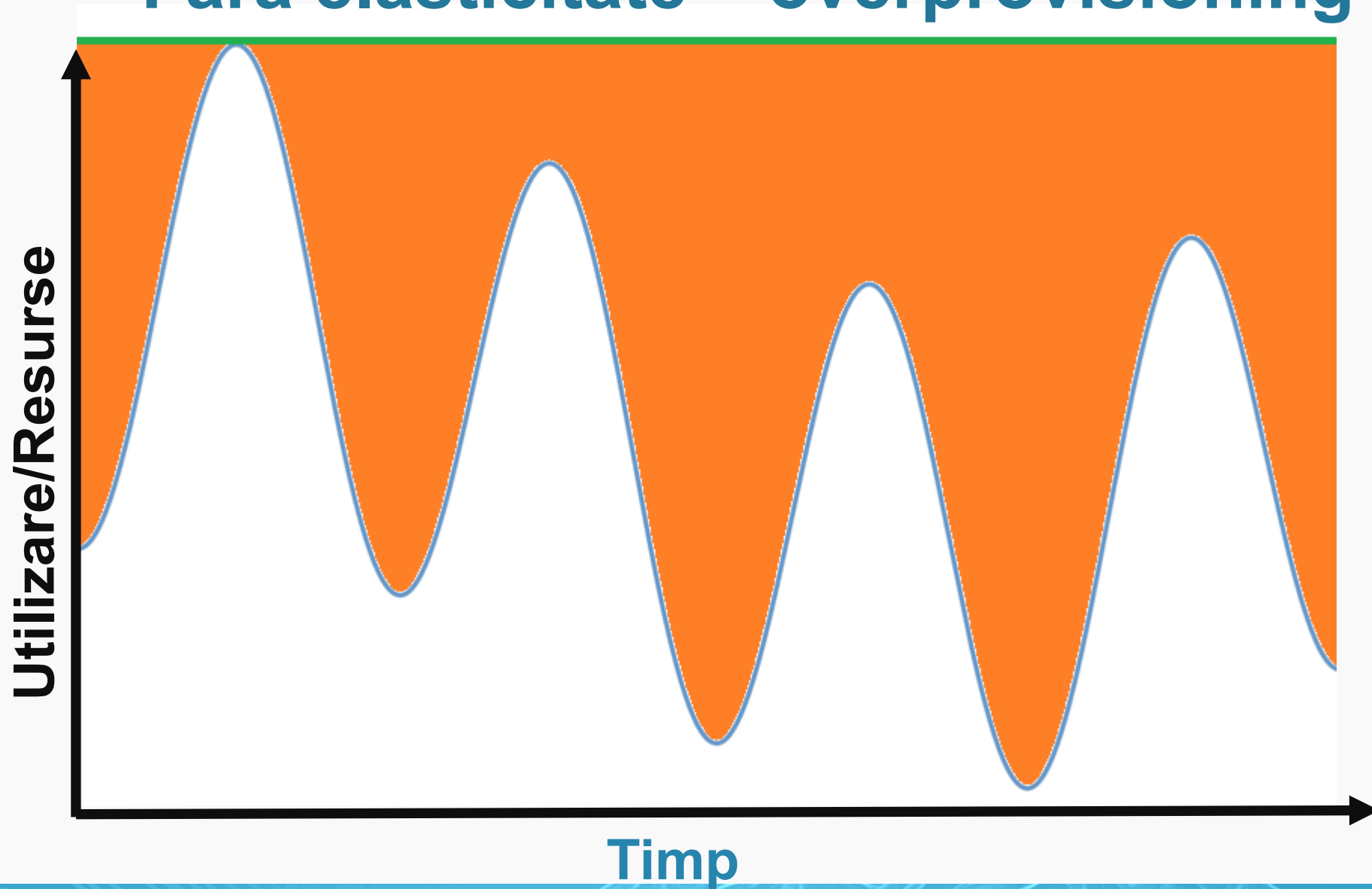


Fără elasticitate – underprovisioning



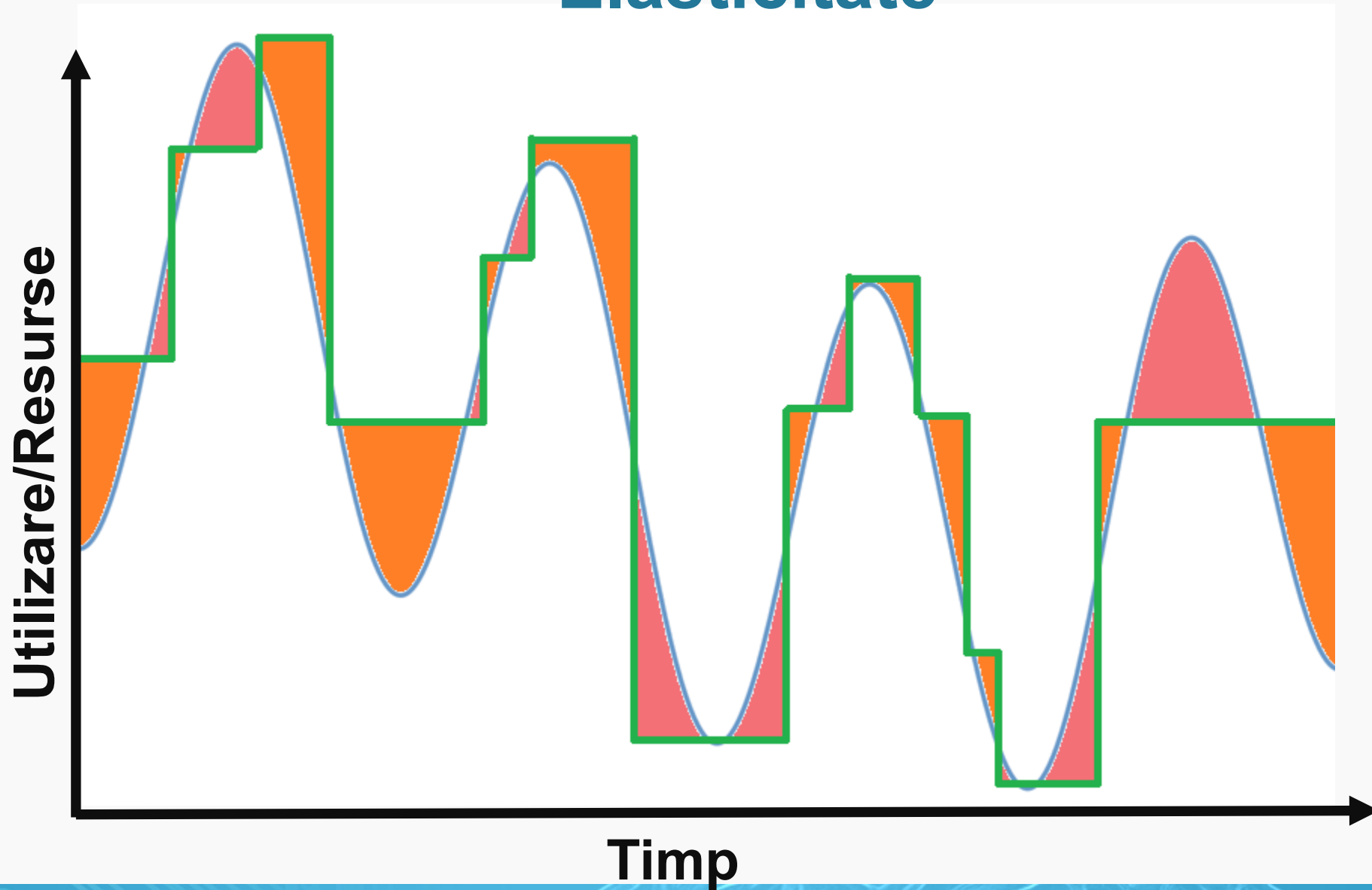


Fără elasticitate – overprovisioning





Elasticitate





Privat vs

Cloud ținut în interiorul unei instituții. Menit pentru a oferi resurse diferitelor grupuri din instituție. Necesită menținerea unui singur grup de administrare cluster.

Public



Google Cloud

Și **mulți** alții.....

vs Hibrid

Și privat și public. Când resursele private sunt epuizate se pot extinde cu cele publice. Posibil doar temporar.



SLA/QoS

Five nines

- Se dorește 99.999% disponibilitate (5 minute pe an down-time)
- Se obține prin securitate crescută
- Se obține prin hardware de înaltă performanță
- Se obține prin redundanță și back-up
- Se obține prin mecanisme de toleranță la defecte



X as a Service

X as a Service	Exemplu Efectiv
Software as a Service	Dropbox, Slack
Platform as a Service	Baze de date
Infrastructure as a Service	Mașini virtuale
Metal as a Service	Mașini fizice
... as a Service	



Metal as a Service

Folosim direct mașinile.

Sunt necesare servere specializate

- Suport **Intelligent Platform Management Interface (IPMI)**
- Suport **Baseboard management controller**
- Suport **Preboot eXecution Environment (PXE)**

Combinate acestea oferă metode de power on/off, , monitorizare, remote-control, instalare OS.

Toate de la distanță





Infrastructure as a Service – mașini virtuale



Hipervizoare:





Avantaje mașini virtuale?



Avantaje mașini virtuale?

Development în sistem de operare diferit

Deployment mai ușor

Dependințe mai ușor de controlat

Scalabilitate mai ușoară (copiezi mașina virtuală)



Emulare vs virtualizare

Simulează sistemul în software

Fiecare instrucțiune e citită de software și interpretată

Ex: [Simulator NES](#) (Warning multiple hours of intense fun)

- Majoritatea instrucțiunilor sunt rulate **DIRECT** de procesor
- **De citit** Capitol 7 – Modern Operating Systems (4th ed) Andrew Tanenbaum



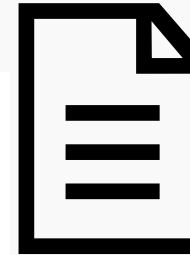
Reminder multi-tasking

Assembly vs. machine code

Machine code bytes	Assembly language statements
	foo:
B8 22 11 00 FF	movl \$0xFF001122, %eax
01 CA	addl %ecx, %edx
31 F6	xorl %esi, %esi
53	pushl %ebx
8B 5C 24 04	movl 4(%esp), %ebx
8D 34 48	leal (%eax,%ecx,2), %esi
39 C3	cmpl %eax, %ebx
72 EB	jnae foo
C3	retl

Instruction stream

```
B8 22 11 00 FF 01 CA 31 F6 53 8B 5C 24
04 8D 34 48 39 C3 72 EB C3
```



Codul de rulat

Loader
(Kernel
Sistemul de
operare)



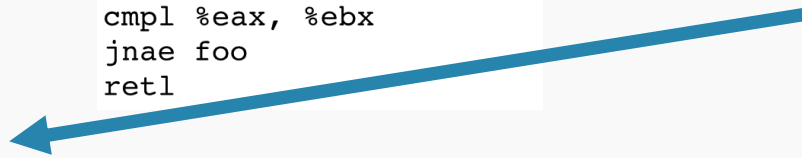


Reminder multi-tasking

Program counter



```
movl $0xFF001122, %eax
addl %ecx, %edx
xorl %esi, %esi
pushl %ebx
movl 4(%esp), %ebx
leal (%eax,%ecx,2), %esi
cmpl %eax, %ebx
jnae foo
retl
```





Reminder multi-tasking

Program counter



```
movl $0xFF001122, %eax
addl %ecx, %edx
xorl %esi, %esi
pushl %ebx
movl 4(%esp), %ebx
leal (%eax,%ecx,2), %esi
cmpl %eax, %ebx
jnae foo
retl
```





Reminder multi-tasking

Program counter



```
movl $0xFF001122, %eax
addl %ecx, %edx
xorl %esi, %esi
pushl %ebx
movl 4(%esp), %ebx
leal (%eax,%ecx,2), %esi
cmpl %eax, %ebx
jnae foo
retl
```





Reminder multi-tasking

Program counter



```
movl $0xFF001122, %eax
addl %ecx, %edx
xorl %esi, %esi
pushl %ebx
movl 4(%esp), %ebx
leal (%eax,%ecx,2), %esi
cmpl %eax, %ebx
jnae foo
retl
```



Timer Interrupt





Reminder multi-tasking

Program counter



```
movl $0xFF001122, %eax
addl %ecx, %edx
xorl %esi, %esi
pushl %ebx
movl 4(%esp), %ebx
leal (%eax,%ecx,2), %esi
cmpl %eax, %ebx
jnae foo
retl
```



Pornește **scheduler**
(posibil read din RAM
pentru cod de scheduler)

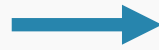
Timer Interrupt





Reminder multi-tasking

Program counter



```
movl $0xFF001122, %eax
addl %ecx, %edx
xorl %esi, %esi
pushl %ebx
movl 4(%esp), %ebx
leal (%eax,%ecx,2), %esi
cmpl %eax, %ebx
jnae foo
retl
```



Scheduler – kernel OS





Reminder multi-tasking



Scheduler – kernel OS

Program counter
State
Registers
Codul e deja acolo



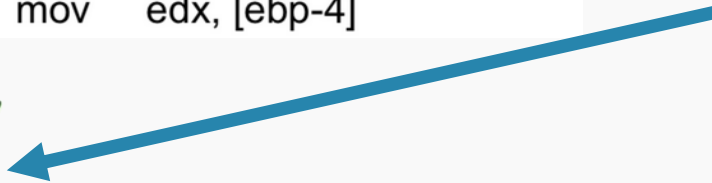


Reminder multi-tasking

```
mov    eax, [ebp+0Ch]
mov    ecx, [eax+4]
push   ecx
call   dword ptr ds:__imp__atoi
add    esp, 4
mov    [ebp-4], eax
mov    edx, [ebp-4]
```



Scheduler – kernel OS





Reminder multi-tasking



```
mov    eax, [ebp+0Ch]
mov    ecx, [eax+4]
push   ecx
call   dword ptr ds:__imp__atoi
add    esp, 4
mov    [ebp-4], eax
mov    edx, [ebp-4]
```



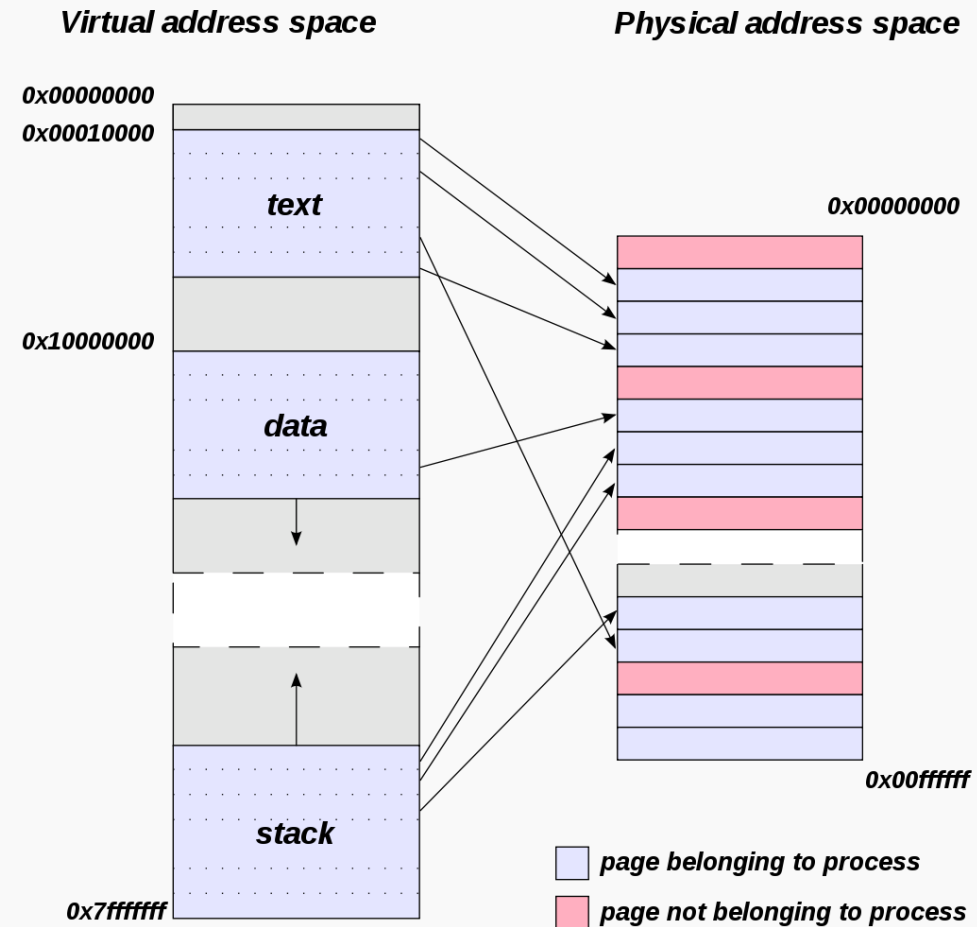


Reminder multi-tasking

Toate programele cred că au acces la tot RAM-ul fără spațiu alocat kernel

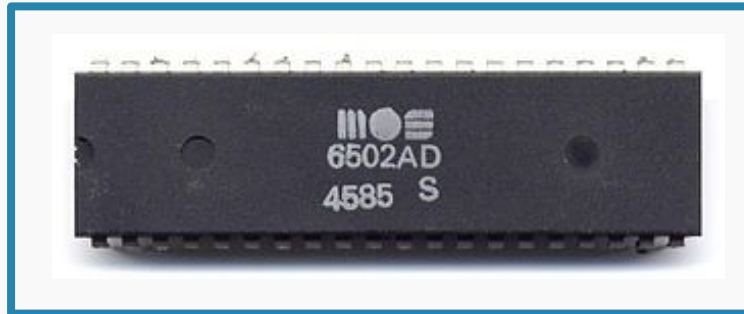
Soluția este paginarea (bucăți de RAM virtual mapate pe cel real)

Kernel-ul și CPU (prin MMU - memory management unit și TLB - translation lookaside buffer) controlează RAM-ul prin loader, malloc, free și tabelele de pagini





Emulare CPU-RAM



Trebuie reținută starea (un set de variabile)
Trebuie reținuți regiștrii (un vector)
Majoritatea instrucțiunilor modifică starea sau regiștrii
Restul instrucțiunilor sunt citiri/scrieri RAM sau periferice



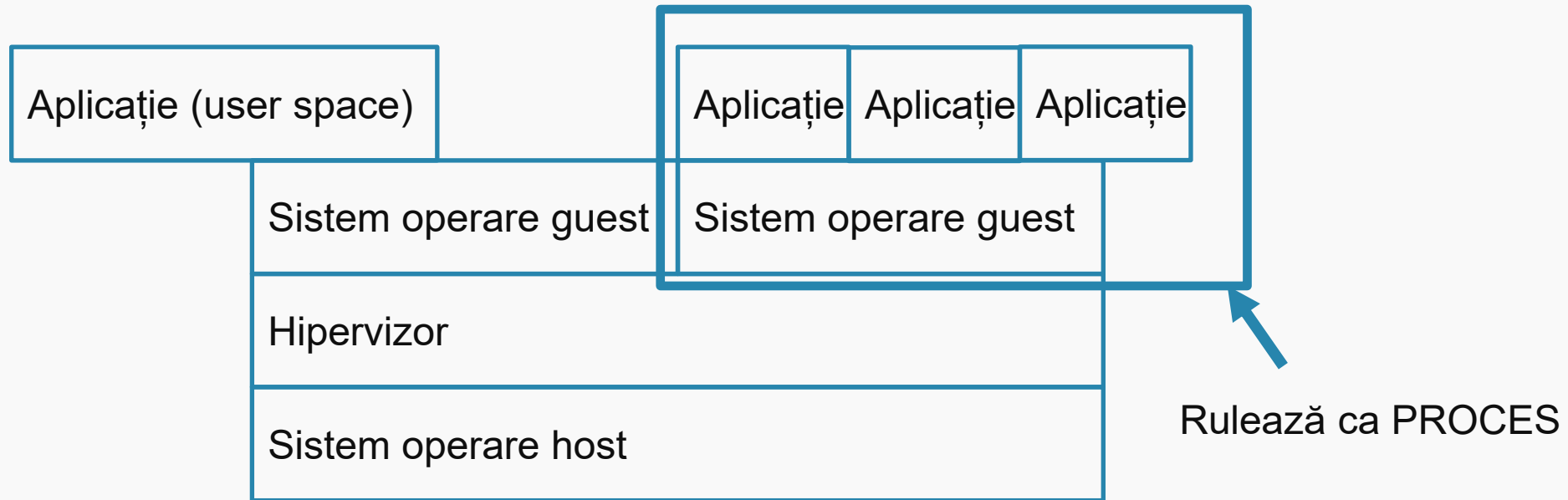
Codul de rulat

Un malloc mare





Virtualizare CPU





Virtualizare CPU

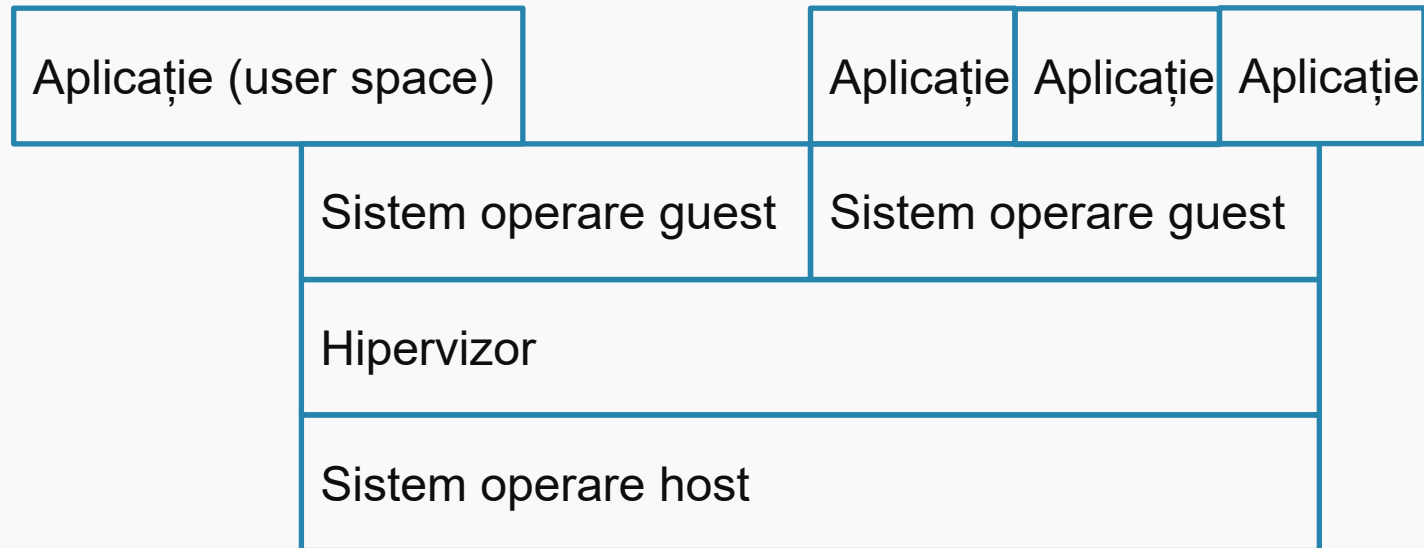
Formal Requirements for Virtualizable Third Generation Architectures

Gerald J. Popek
University of California, Los Angeles
and
Robert P. Goldberg
Honeywell Information Systems and
Harvard University

Virtual machine systems have been implemented on a limited number of third generation computer systems, e.g. CP-67 on the IBM 360/67. From previous empirical studies, it is known that certain third generation computer systems, e.g. the DEC PDP-10, cannot support a virtual machine system. In this paper, model of a third-generation-like computer system is developed. Formal techniques are used to derive precise sufficient conditions to test whether such an architecture can support virtual machines.

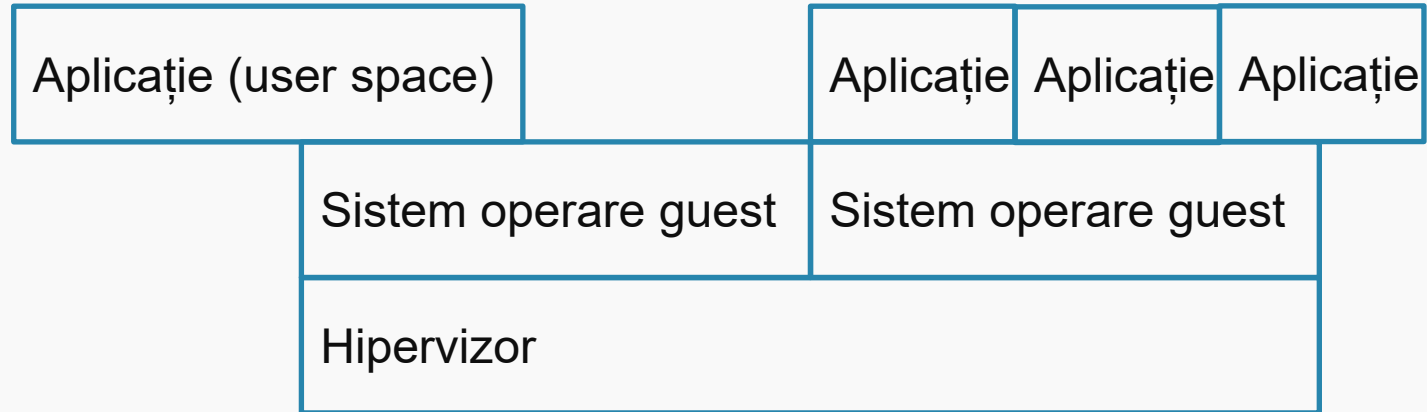


Virtualizare CPU – Type 1



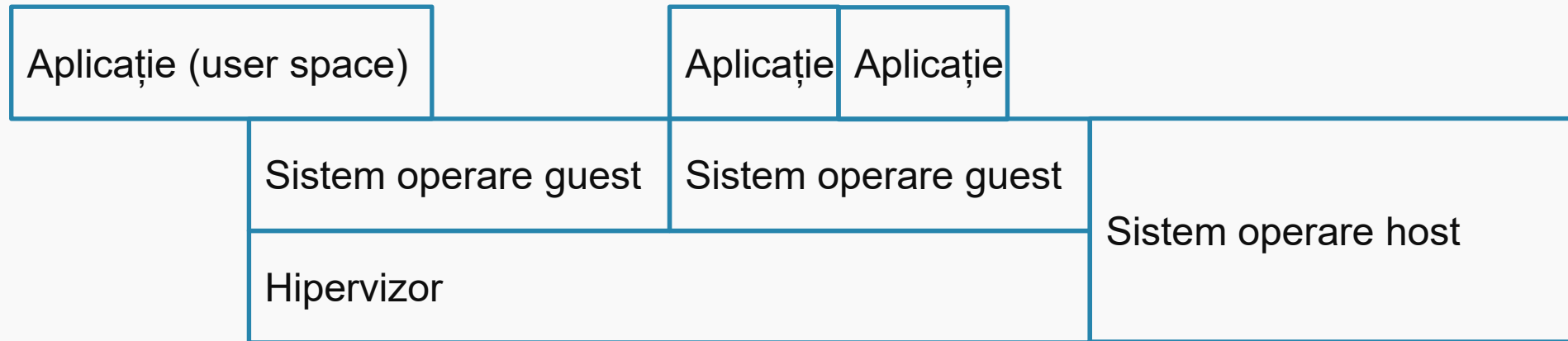


Virtualizare CPU – Type 2





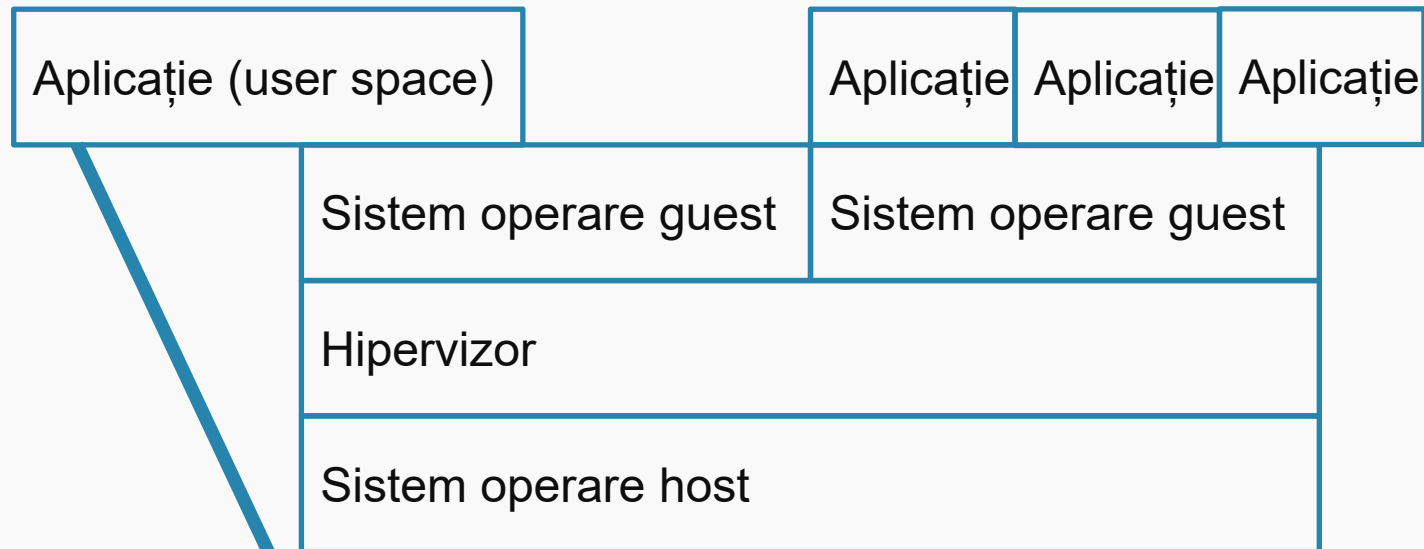
Virtualizare CPU – Paravirtualizare





Virtualizare CPU

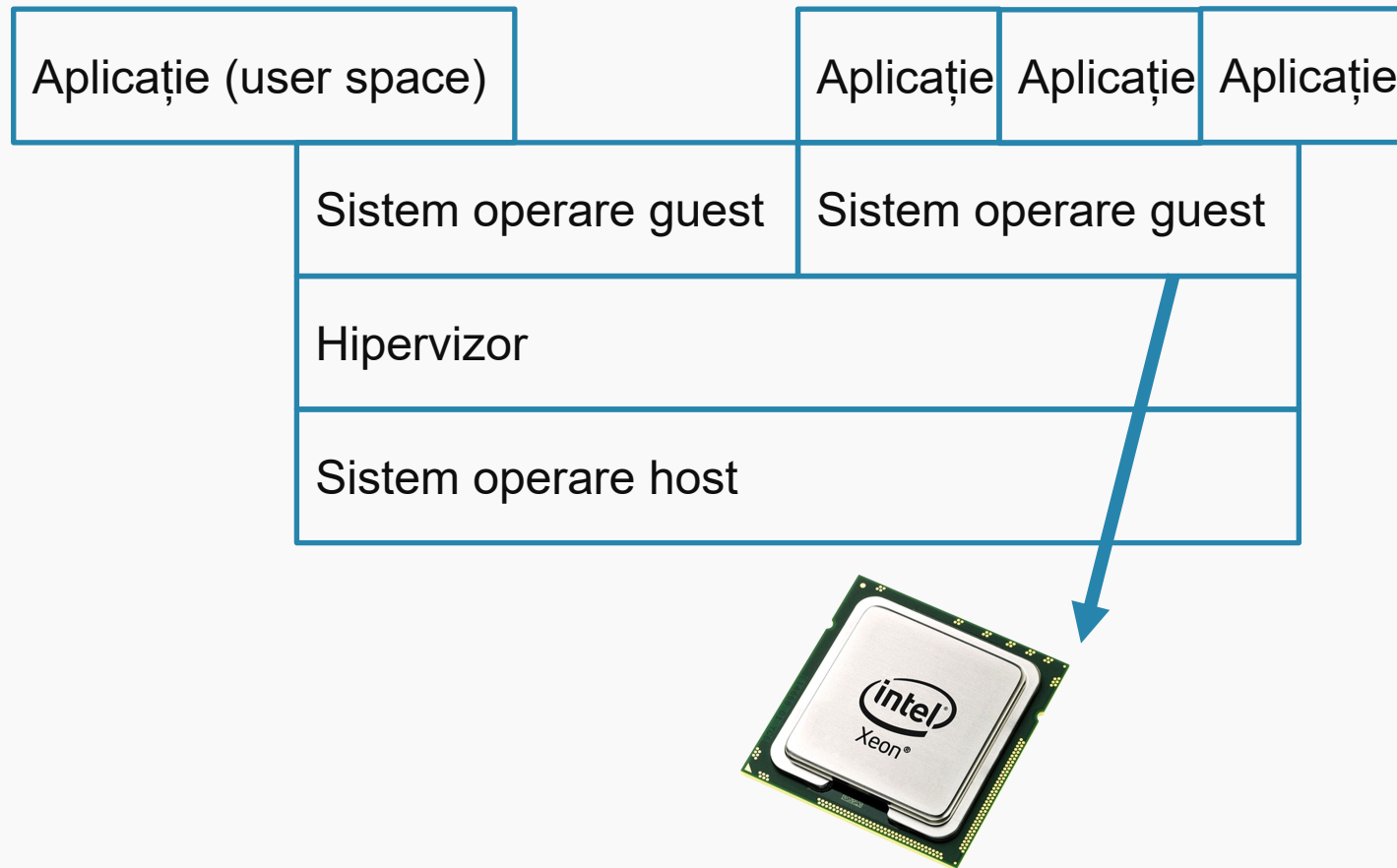
Instrucțiuni simple (ex. add, mov din memoria sa)





Virtualizare CPU – trap and execute

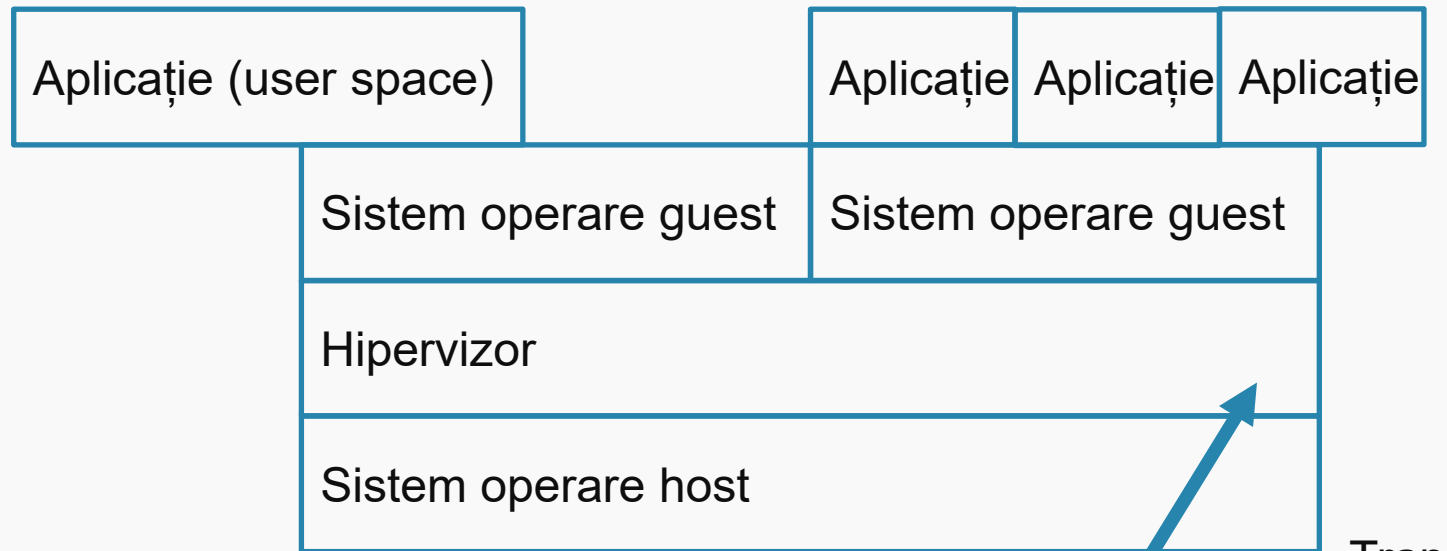
Instrucțiuni kernel (ex. umblat cu I/O, sist întreruperi)





Virtualizare CPU – trap and execute

Instrucțiuni kernel (ex. umblat cu I/O, sist întreruperi)



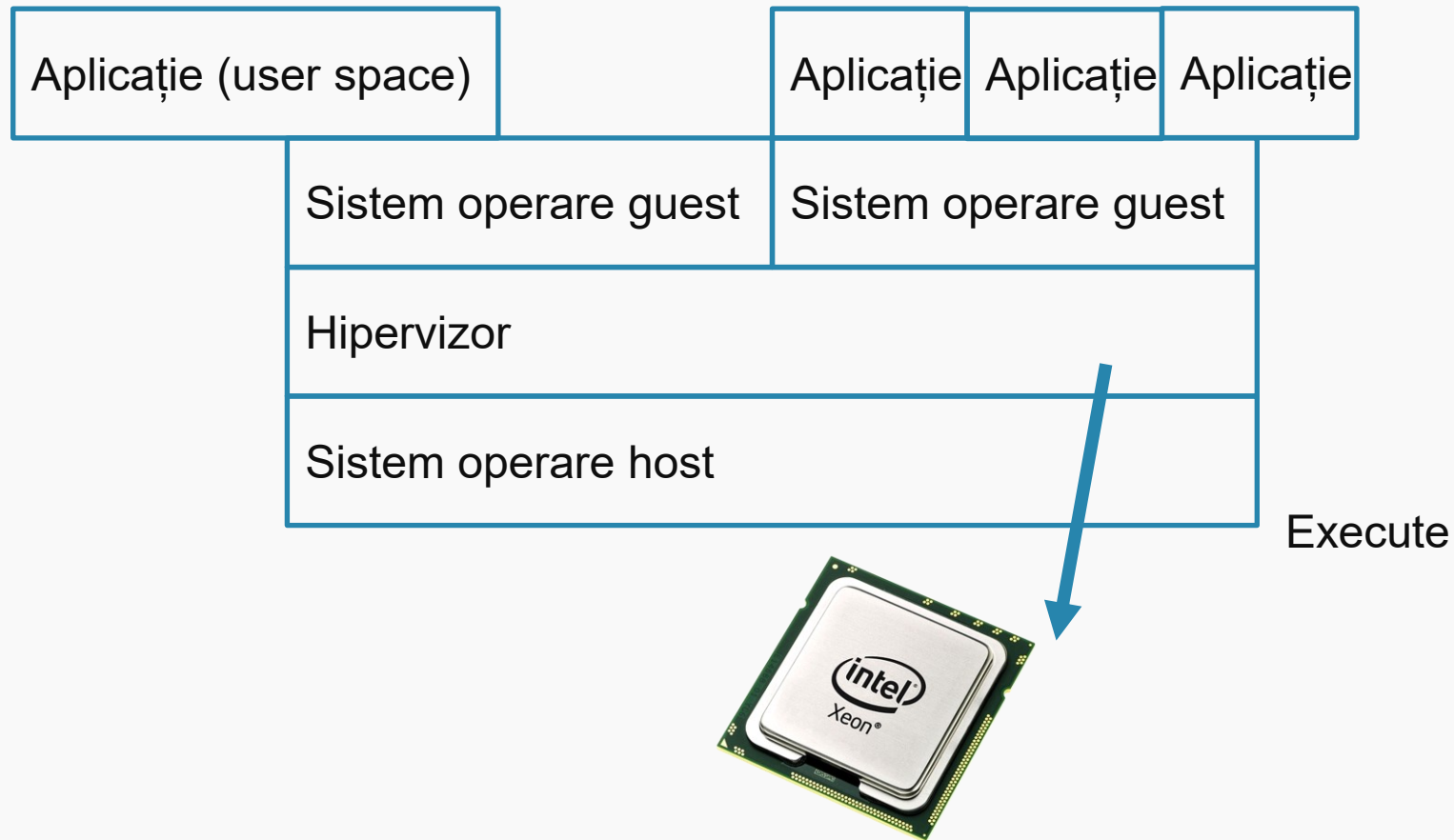
Trap





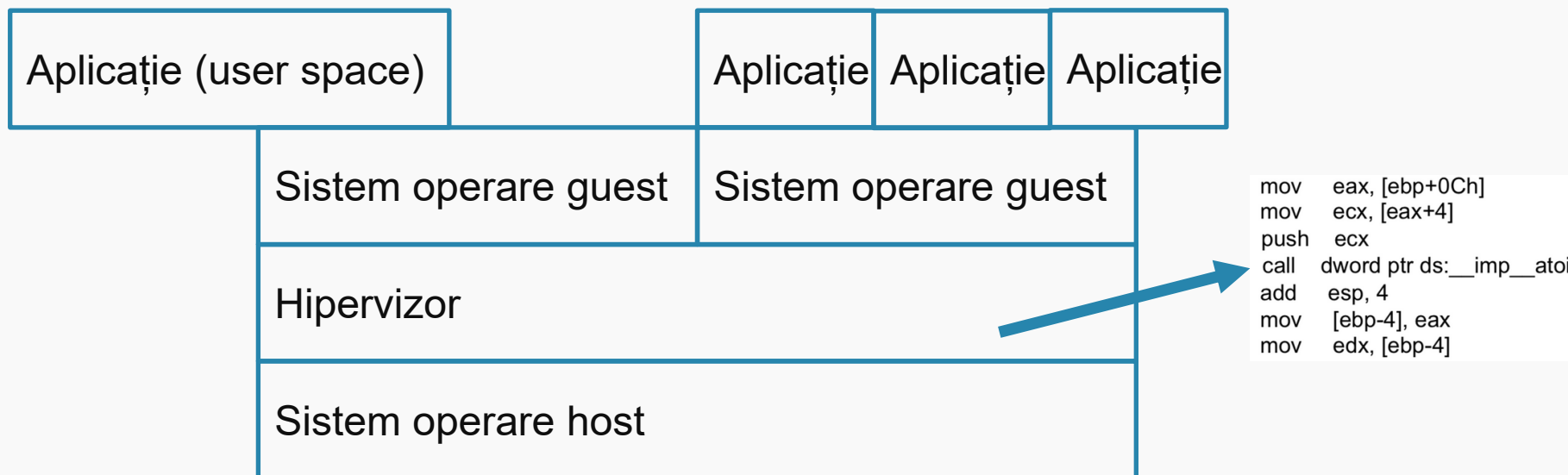
Virtualizare CPU – trap and execute

Instrucțiuni kernel (ex. umblat cu I/O, sist întreruperi)





Virtualizare CPU – rescriere instrucțiuni (emulare) Instrucțiuni kernel (ex. umblat cu I/O, sist întreruperi)



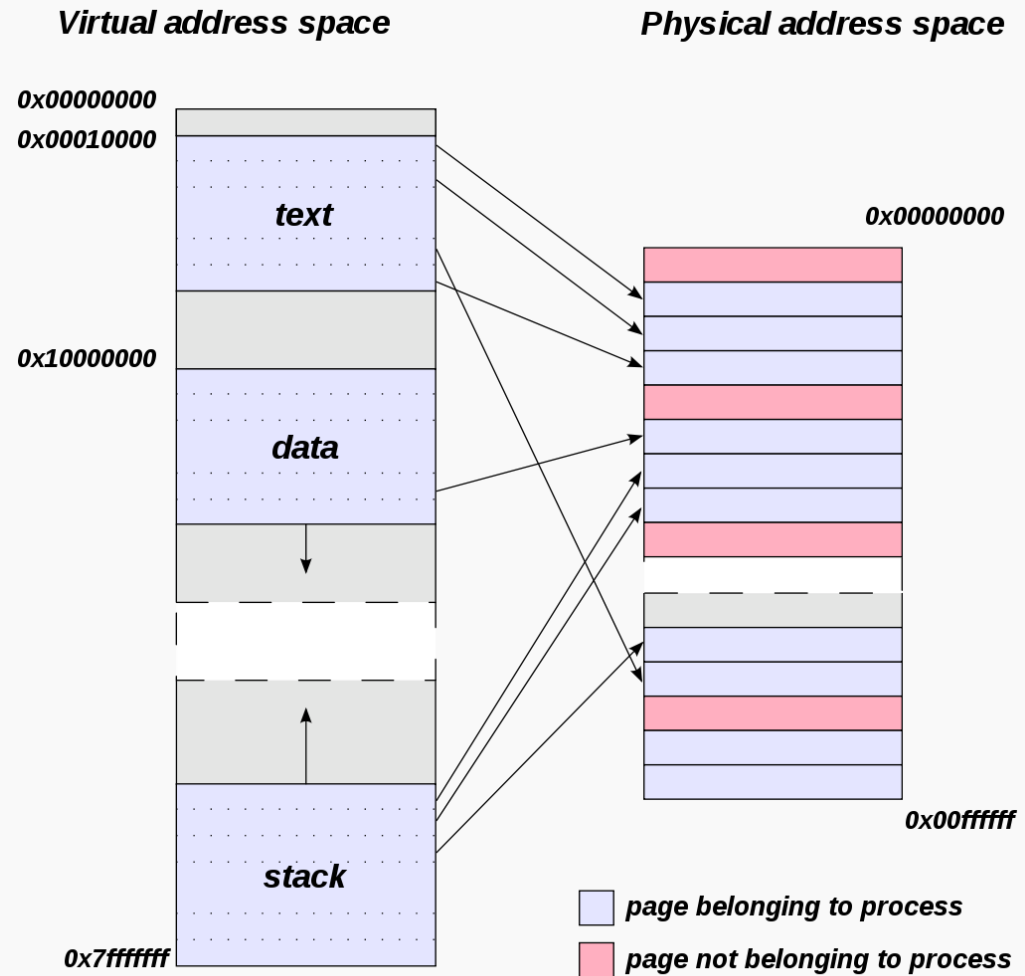
Înainte să ajungă să fie executate
instrucțiunile sunt modificate (per
bloc de instrucțiuni)
Inclusiv branch e modificat



Virtualizare RAM

Shadow page tables

Suport direct în hardware



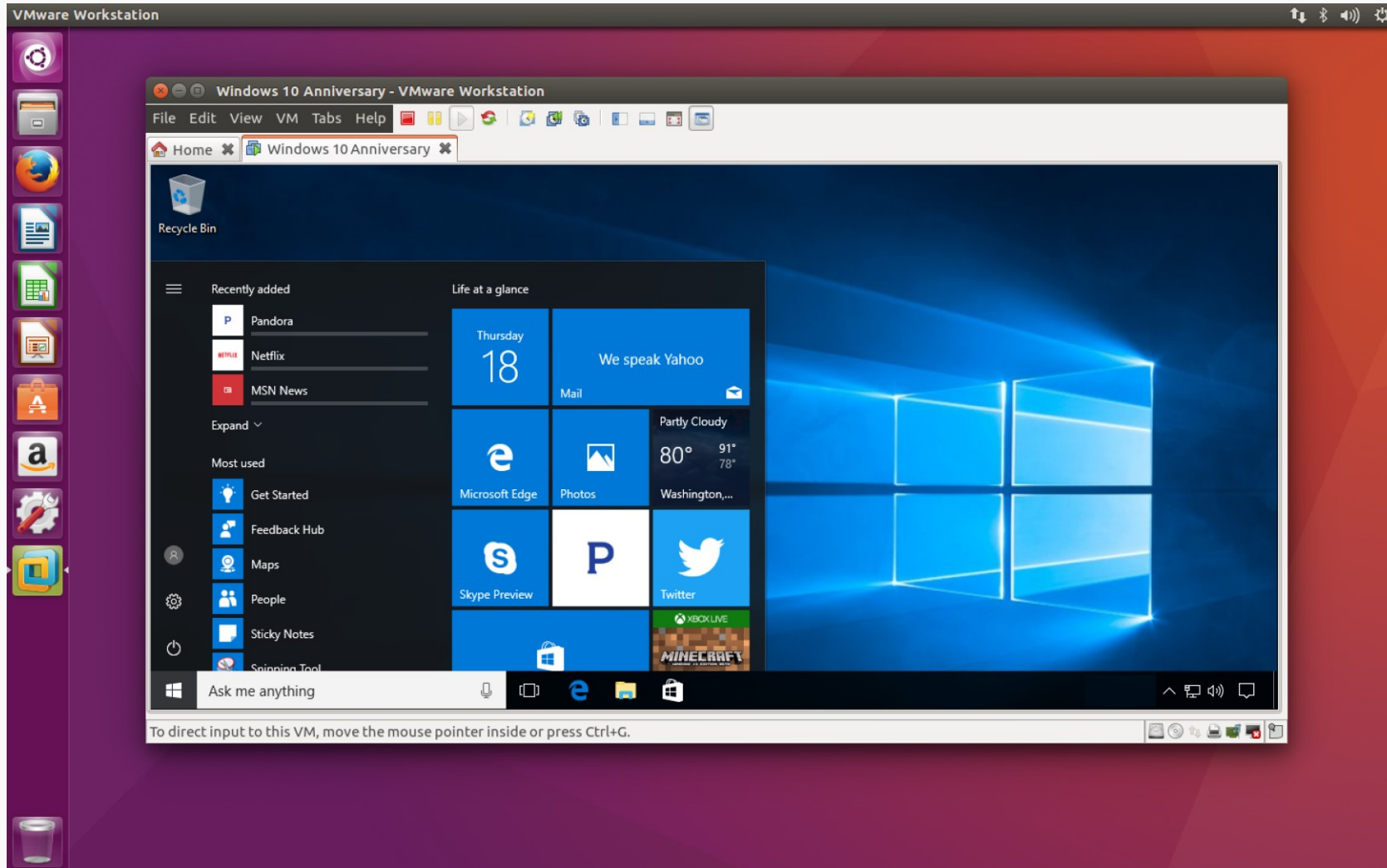


Virtualizare Emulare HDD





Virtualizare Emulare Display





Virtualizare Emulare I/O

Mouse/Tastatură ?



Virtualizare Emulare I/O

Mouse/Tastatură ?

Idem Display

Multe avantaje când vine vorba de drivere

- Dacă există sistem de operare host se pot folosi driverele acestuia pentru periferice
- Pentru sistemul de operare guest se poate oferi un singur dispozitiv virtual (un singur driver)
- Portabilitate – se poate modifica hardware fără să modifice mașina virtuală



Platform as a Service

Cea mai mare dezvoltare în ultimii ani.

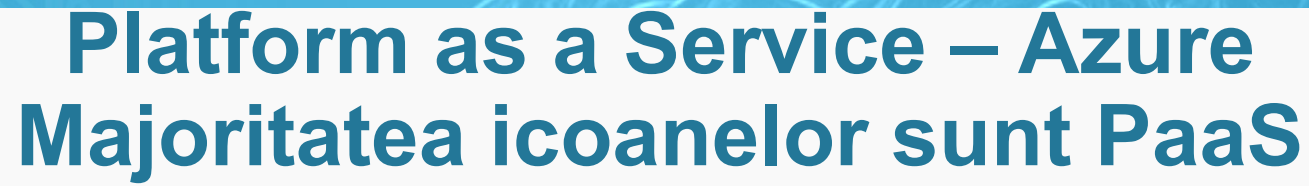
Baze de date cu capabilități diverse (sql/no-sql, diferite nivele de acid, diferite nivele de scalabilitate, diferite nivele de consistență)

Analytics

Blockchain

Inteligență artificială

- Face detection
- Chat bots
- GTP-3

Cristian Chilipirea – Sisteme Tolerante la Defecte



Platform as a Service – GCP (Majoritatea) Uitați-vă la scrollbar

The screenshot displays the Google Cloud Platform (GCP) website. The top navigation bar includes links for 'ricing', 'Getting Started', and 'Contact Us', along with a search icon, 'Docs', 'Support', a language selector set to 'English', a 'Console' button with a user profile icon, and a 'Get started for free' button. The main content area is titled 'Google Cloud products' and features two buttons: 'Get started for free' and 'See pricing'. Below this, the 'Featured products' section is organized into a grid of product cards, each with a title and a brief description. The sidebar on the left lists various Google Cloud products and services, including 'Overview', 'Featured products', 'AI and Machine Learning', 'API Management', 'Compute', 'Containers', 'Data Analytics', 'Databases', 'Developer Tools', 'Healthcare and Life Sciences', 'Hybrid and Multi-cloud', 'Internet of Things (IoT)', 'Management Tools', 'Media and Gaming', 'Migration', 'Networking', 'Operations', 'Security and Identity', 'Serverless Computing', 'Storage', 'More Google Cloud products', 'Product launch stages', and 'Take the next step'. The right side of the page shows a 'More from Google Cloud' section with links to 'Solutions', 'Training', and 'Certification'. A vertical scrollbar is visible on the right edge of the page, indicating the content extends beyond the visible area.

Google Cloud products

Overview

Featured products

AI and Machine Learning

API Management

Compute

Containers

Data Analytics

Databases

Developer Tools

Healthcare and Life Sciences

Hybrid and Multi-cloud

Internet of Things (IoT)

Management Tools

Media and Gaming

Migration

Networking

Operations

Security and Identity

Serverless Computing

Storage

More Google Cloud products

Product launch stages

Take the next step

Google Cloud products

[Get started for free](#) [See pricing](#)

Featured products

- BigQuery**
Data warehouse for business agility and insights.
- Cloud CDN**
Content delivery network for delivering web and video.
- Cloud Functions**
Event-driven compute platform for cloud services and apps.
- Cloud Run**
Fully managed environment for running containerized apps.
- Cloud SDK**
Command-line tools and libraries for Google Cloud.
- Cloud SQL**
Relational database services for MySQL, PostgreSQL, and SQL Server.
- Cloud Storage**
Object storage that's secure, durable, and scalable.
- Compute Engine**
Virtual machines running in Google's data center.
- Dataflow**
Streaming analytics for stream and batch processing.
- Google Kubernetes Engine**
Managed environment for running containerized apps.
- Anthos**
Platform for modernizing existing apps and building new ones.
- Operations**
Monitoring, logging, and application performance suite.

AI and Machine Learning



Platform as a Service – AWS (Majoritatea) Toate se deschid



Analytics



Application Integration



AWS Cost Management



Blockchain



Business Applications



Compute



Containers



Customer Engagement



Database



Developer Tools

Amazon Connect
Cloud-based Contact Center

Amazon Pinpoint
Personalized User Engagement Across Channels

Amazon Simple Email Service (SES)
Email Sending and Receiving



End User Computing



Front-End Web & Mobile



Game Tech



Internet of Things



Machine Learning



Management & Governance



Media Services



Migration & Transfer



Networking & Content
Delivery



Quantum Technologies



Robotics



Satellite



Security, Identity &
Compliance



Serverless



Storage

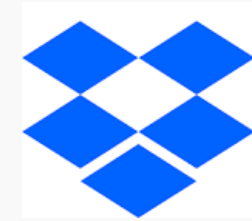


VR & AR

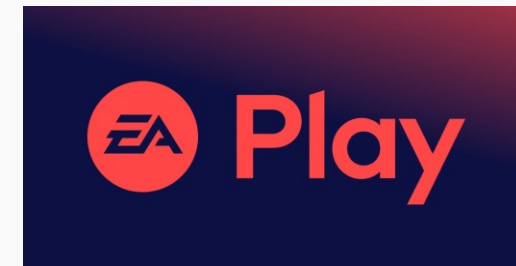


Software as a Service

Majoritatea companiilor merg în această direcție



NETFLIX





Azure for students

<https://azure.microsoft.com/en-us/free/students/>

Start building the future with Azure for Students!

Get a \$100 credit when you create your free Azure for Students account

[Activate now](#)

[Read the FAQ for eligibility >](#)

No credit card needed

Simply verify your student status through your school email address, and you're ready to go with \$100 in credit.

+

Free developer tools

Build your skills in trending tech including data science, artificial intelligence (AI), machine learning, and other areas with access to professional developer tools.

[Learn more >](#)

+

Drive your career

Access free learning paths and labs to take your skills to the next level.



Azure for students

Folosiți contul de email mta.ro

Parola este cea de pe email mta.ro (posibil diferită de cea de pe wiki)

După intrarea în cont, pe Azure trebuie activați cei 100\$.