POLITECNICO DI MILANO

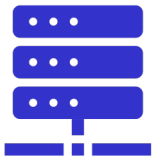
Computing Infrastructures





Software Infrastructures: Cloud Computing and Edge/Fog Computing

The topics of the course



A. HW Infrastructures:

- **System-level**: Computing Infrastructures and Data Center Architectures, Rack/Structure;
- **Node-level**: Server (computation, HW accelerators), Storage (Type, technology), Networking (architecture and technology)
- **Building-level**: Cooling systems, power supply, failure recovery

B. SW Infrastructures:

- Computing Architectures: Cloud Computing (types, characteristics), X-as-a service, Edge/Fog Computing
- **Virtualization**: Process/System VM, Virtualization Mechanisms (Hypervisor, Para/Full virtualization)
- **Machine and deep learning-as-a-service**

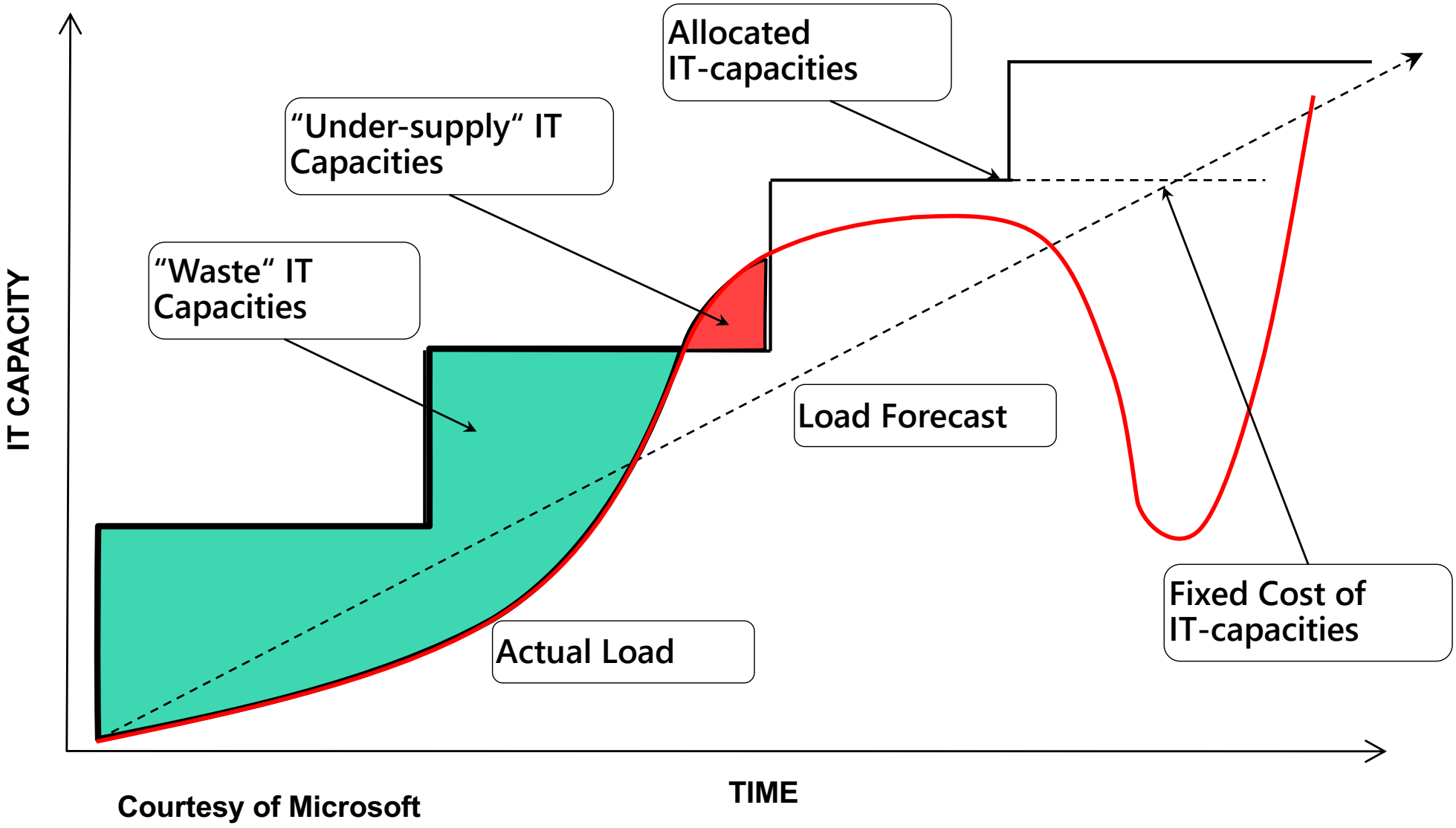
C. Methods:

- **Reliability and availability of datacenters** (definition, fundamental laws, RBDs)
- **Disk performance** (Type, Performance, RAID)
- **Scalability and performance of datacenters** (definitions, fundamental laws, queuing network theory)

What is Cloud Computing?

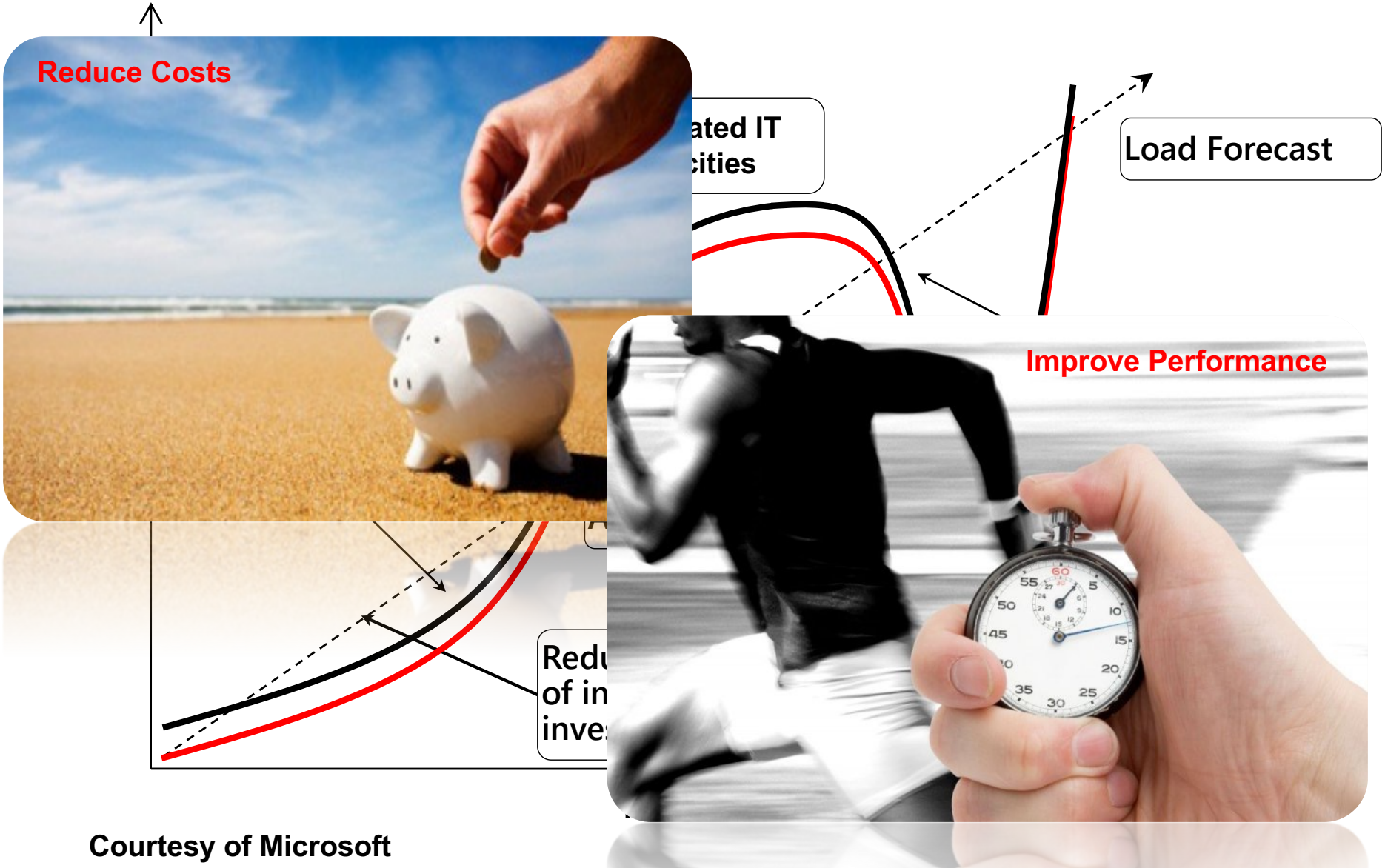
- A coherent, large-scale, publicly accessible collection of computing, storage, and networking resources
- Available via Web service calls through the Internet
- Short- or long-term access on a pay-per-use basis

Over-provisioning - Out of Cloud



Courtesy of Microsoft

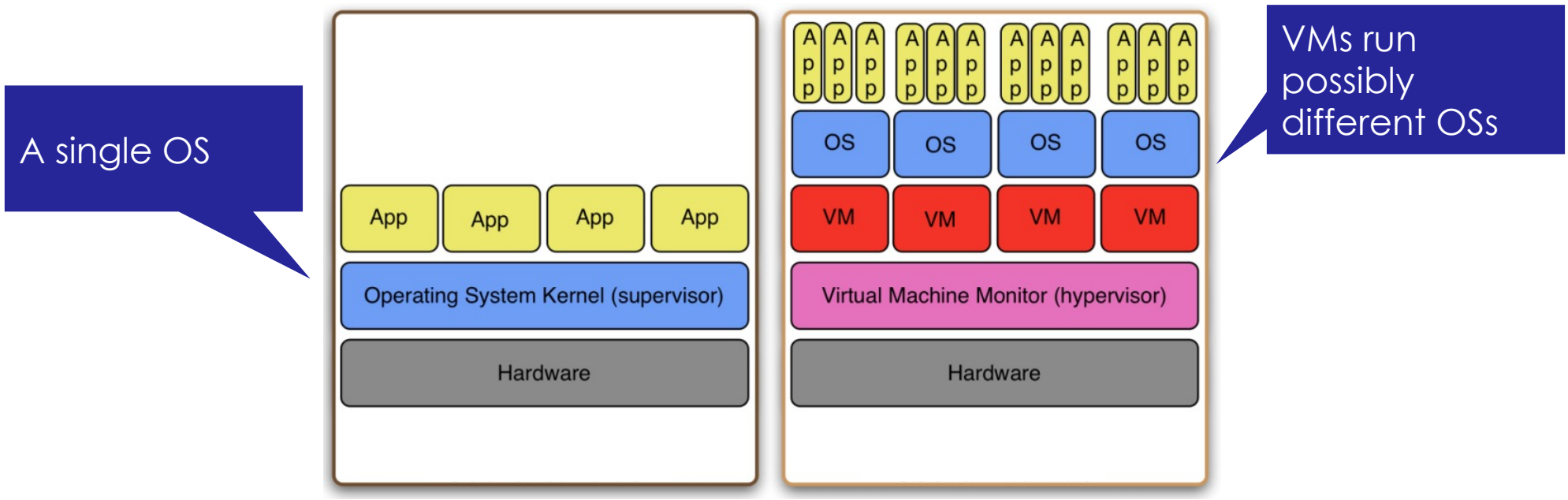
Cloud-provisioning



Courtesy of Microsoft

How is Cloud implemented? Virtualization

- Hardware resources (CPU, RAM, ecc...) are partitioned and shared among multiple **virtual machines** (VMs)
- The virtual machine monitor (VMM) governs the access to the physical resources among running VMs
- Performance isolation and security



Virtualization Consequences

Without virtualization:

- Software strongly linked/related with hardware
  - Move/change an application not an easy task
- To isolate failure/crash the classical model is:
  - 1 server
  - 1 operating system (OS)
  - 1 application, with a resulting low CPU utilization (10-15%)
- Low flexibility

With Virtualization:

- Hw-independence: software/hardware no longer strongly related
- High fexibility thanks to pre-built VMs
- OS and applications can be handled as a «*single entity*»

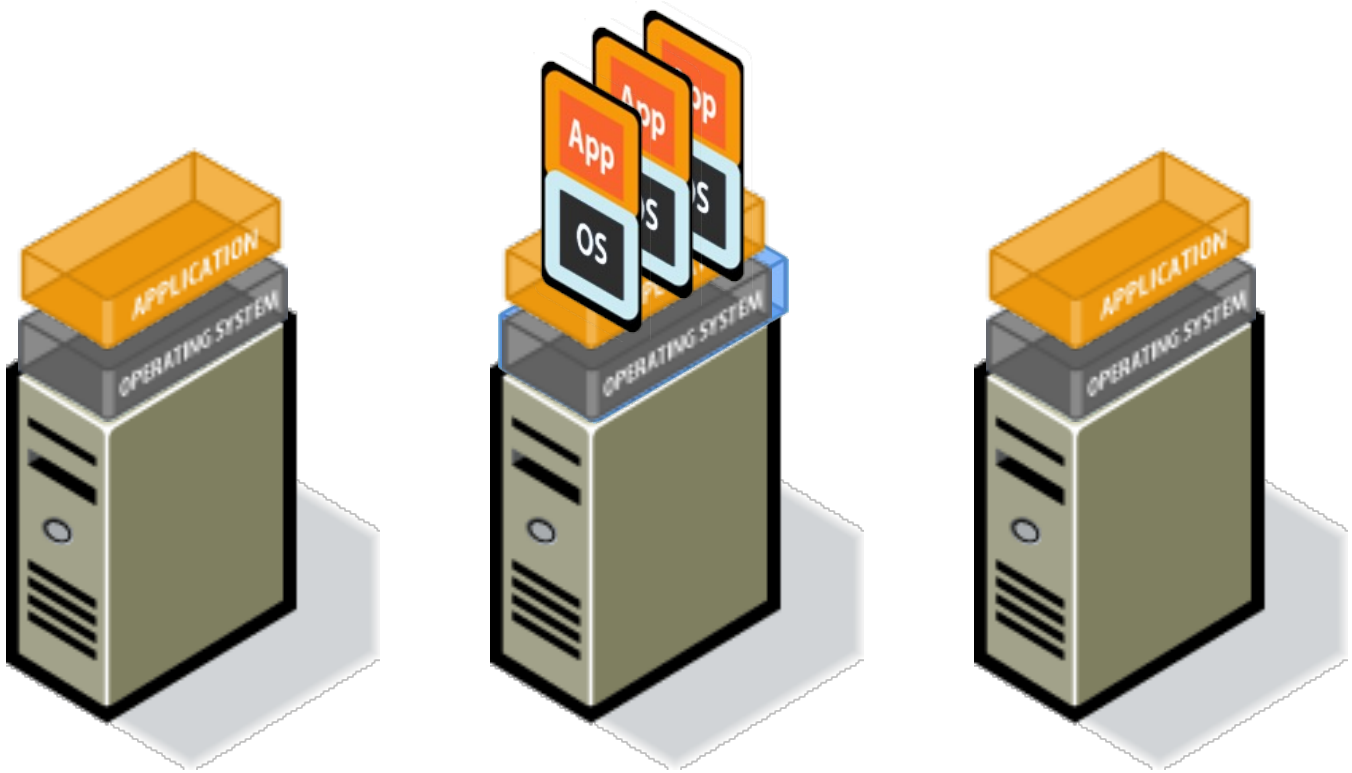
Virtualization Consequences

Impact of Virtualization on the evolution of IT systems:

- Sever consolidation
- Cloud computing

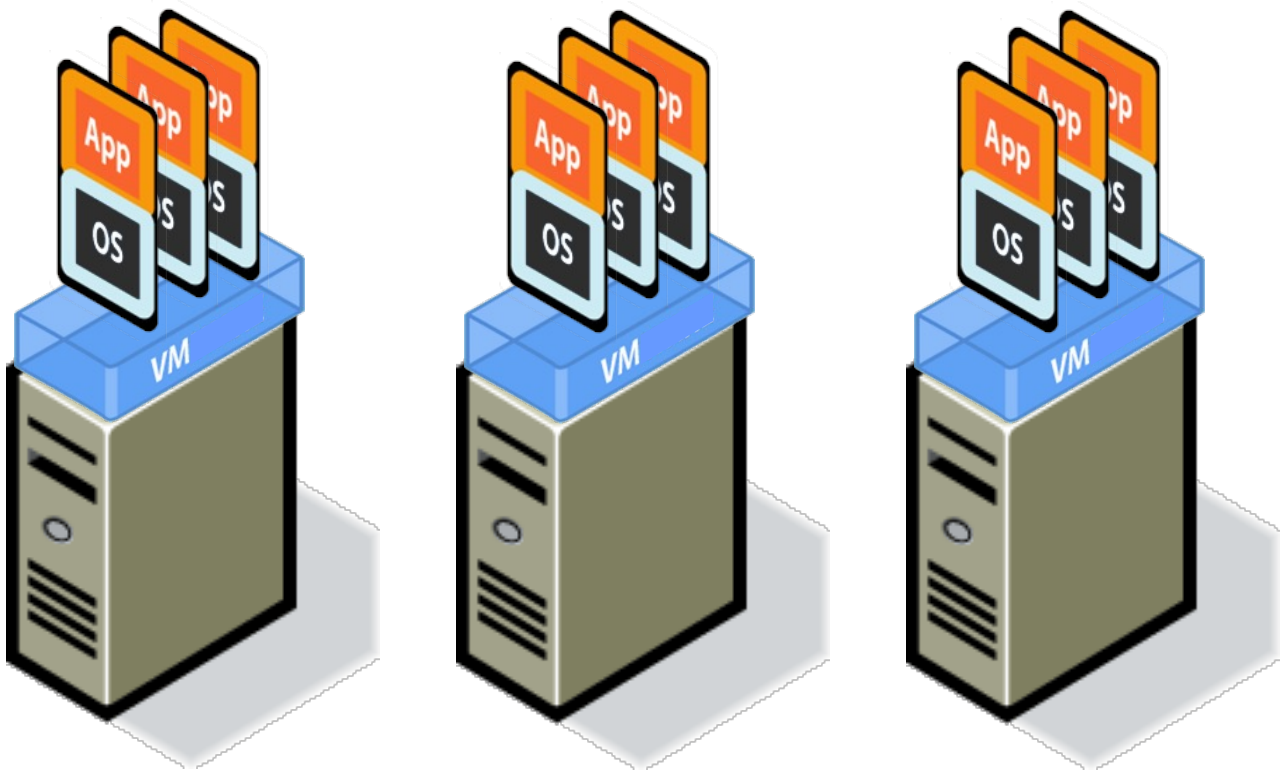
Server Consolidation

Virtualization - Server Consolidation

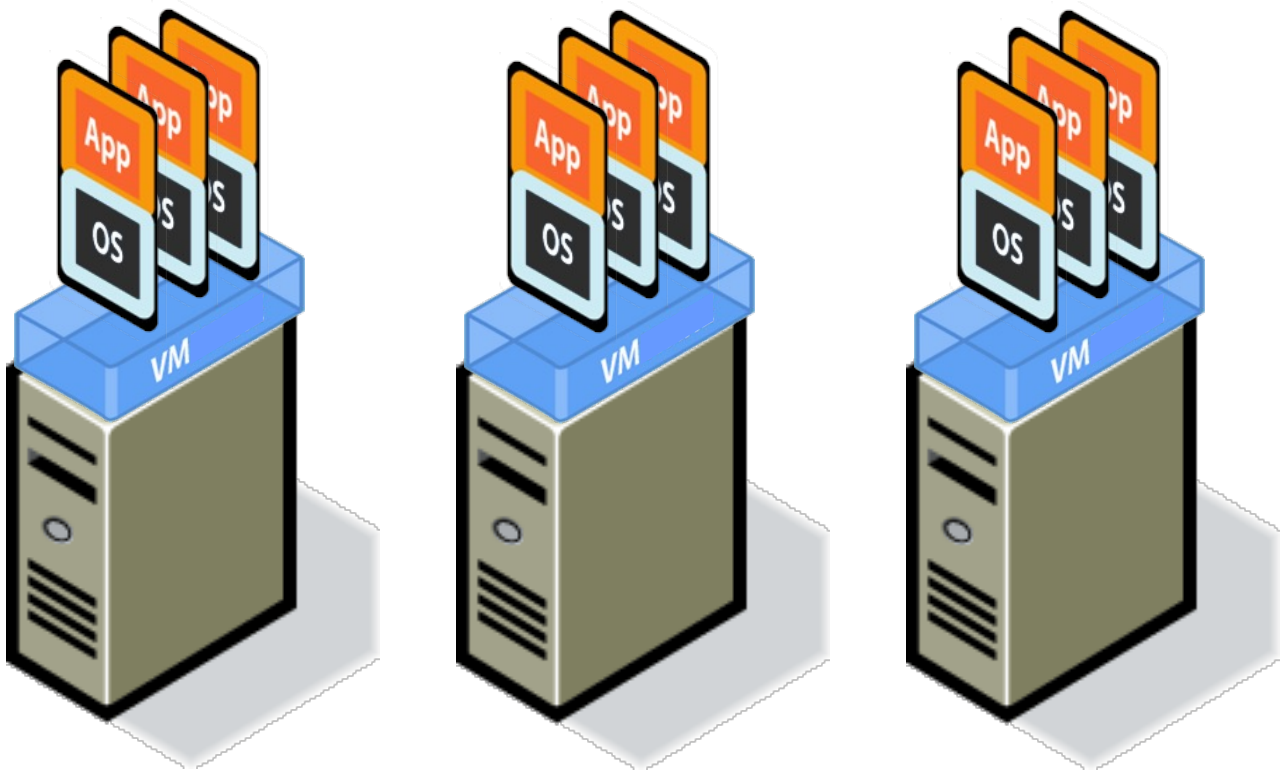


Animation source: VMWare website.

Virtualization

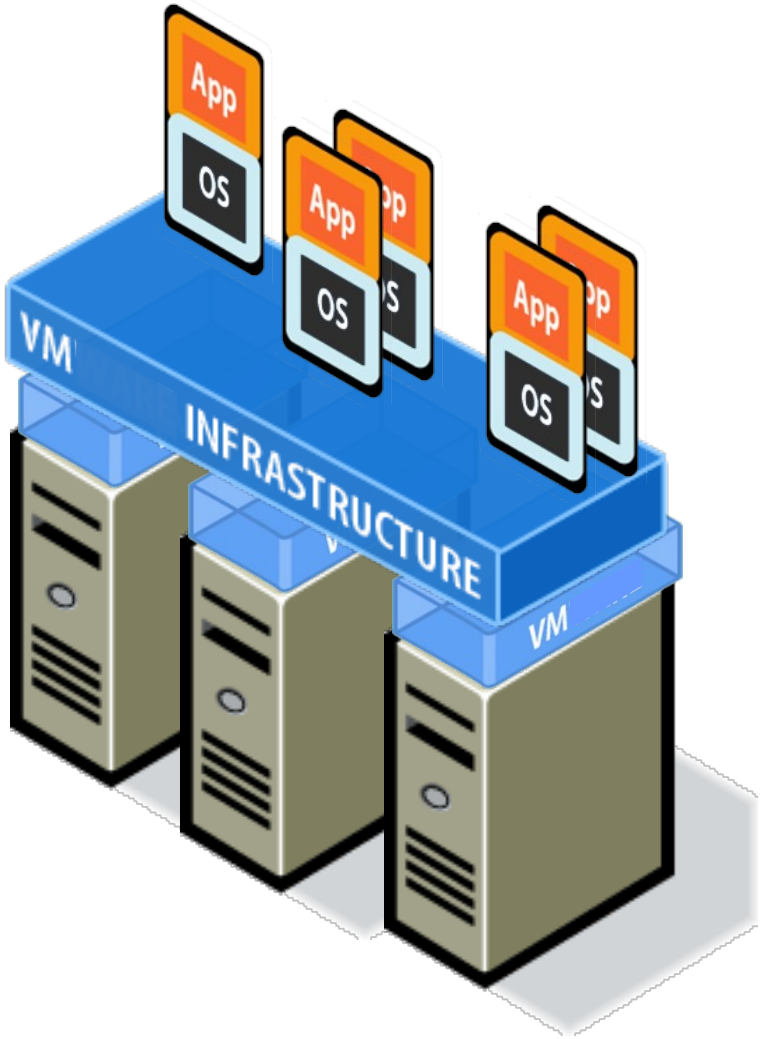


Virtualization



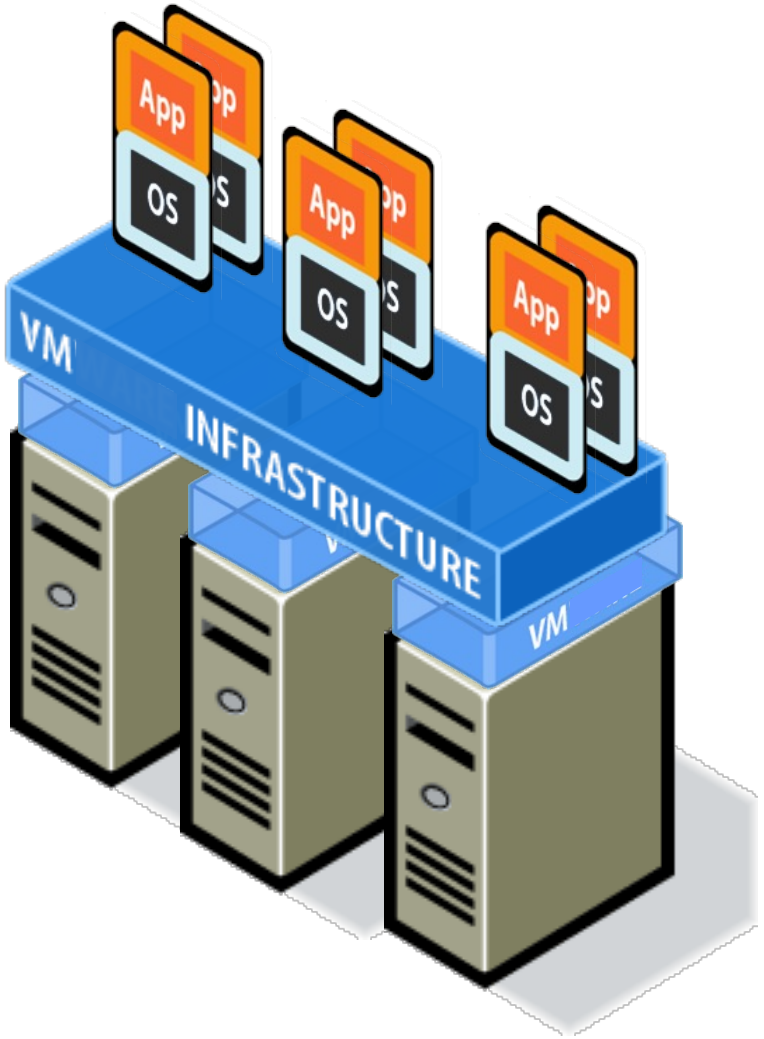
Migration from Physical to Virtual

Consolidation Management: migration from physical to virtual machines



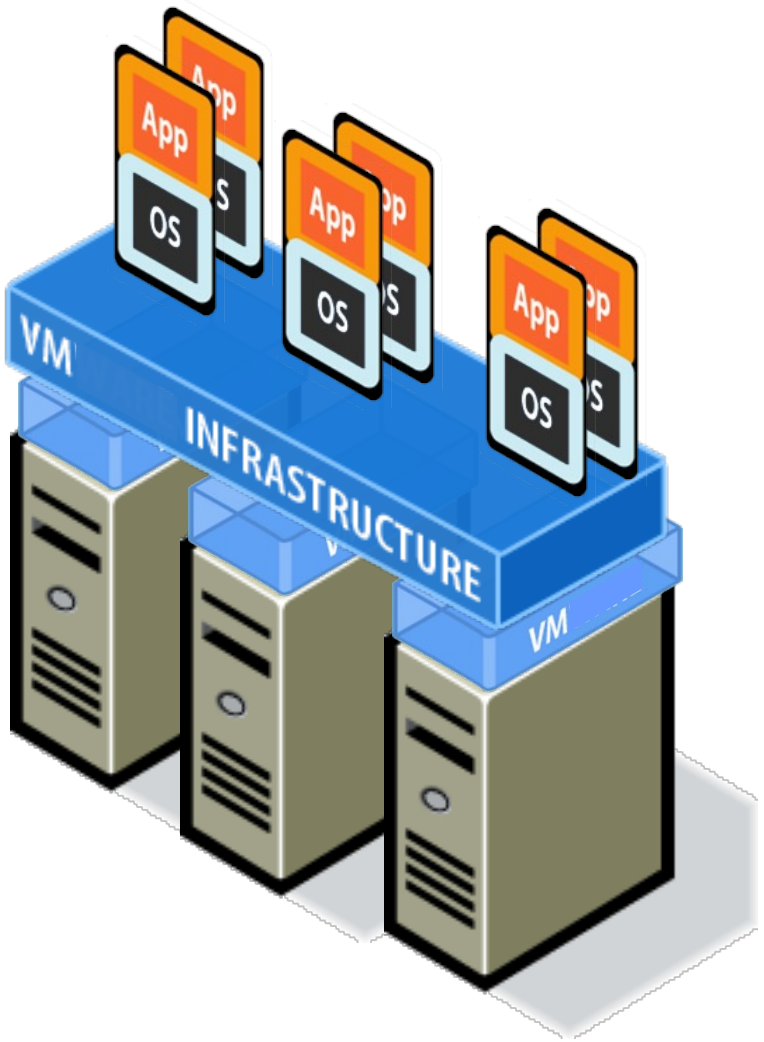
Scalability

It is possible to move Virtual Machines, without interrupting the applications running inside

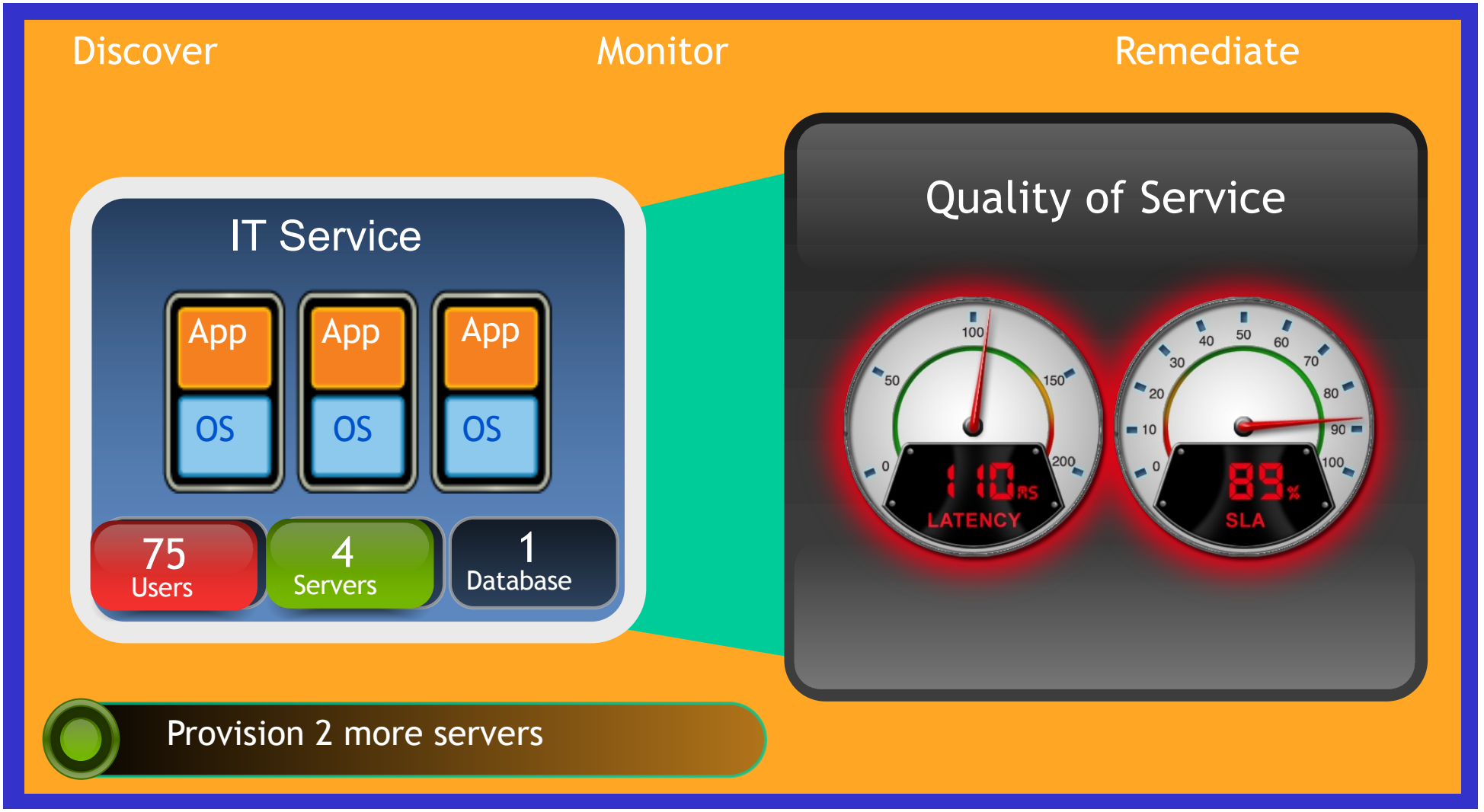


Automatic Scalability

It is possible to automatically balance the Workloads according to set limits and guarantees



Automatic Scalability



High Availability

Servers and Applications are protected against component and system failure

Advantages of consolidation

Consolidation

- Different OS can run on the same hardware
- Higher hardware utilization
  - Less hardware is needed
    - Acquisition costs
    - Management costs (human resources, power, cooling)
  - Green IT-oriented
- Continue to use legacy software (e.g., software for WIN on Linux machines thanks to VMs)
- Application independent from the hardware

Cloud Computing

Cloud Computing: resources as utilities

Cloud computing is a **model for** enabling

- convenient
- **on-demand**

network access to a shared pool of configurable computing resources, like for example:

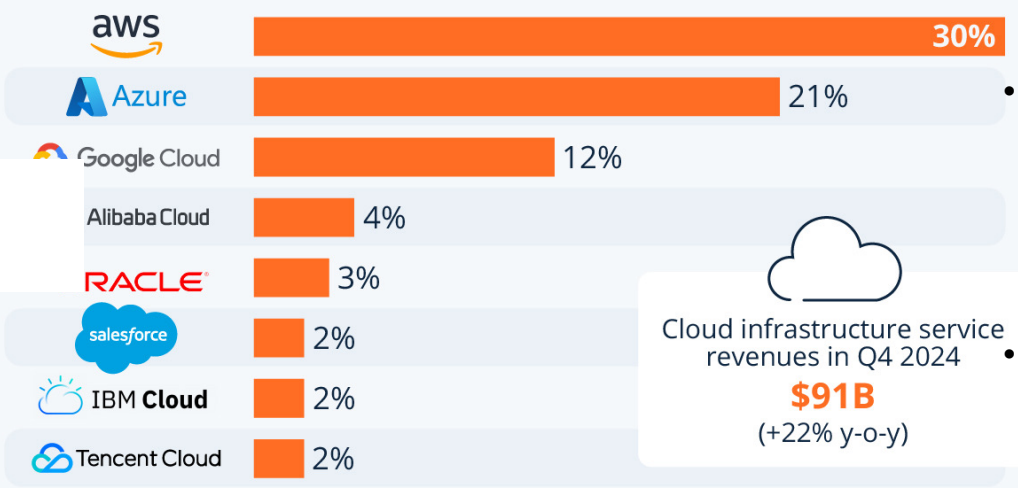
- Networks
- Servers
- Storage
- Applications
- Services

that can be **rapidly provisioned and released with minimal management** effort or service provider interaction

Cloud Computing Growth

Amazon and Microsoft Stay Ahead in Global Cloud Market

Worldwide market share of leading cloud infrastructure service providers in Q4 2024*



Cloud infrastructure service revenues in Q4 2024
$91B
(+22% y-o-y)

* Includes platform as a service (PaaS) and infrastructure as a service (IaaS) as well as hosted private cloud services

Source: Synergy Research Group

statista

- **Compound Annual Growth Rate (CAGR) 22%**
- **$330 billion** market: Despite its size, the cloud market is still growing strongly
- **Increased emphasis on multi-cloud strategy**: According to Accenture, **93% of enterprises** have built up to a **multi-cloud** strategy
- **Increase adoption of hybrid cloud services**: Enterprises having their existing infrastructure are moving toward the adoption of cloud computing services and are willing to adopt the hybrid approach (about **87%** of enterprises have already adopted **hybrid cloud** strategies)
- **Boosting the adoption of edge computing technology**: Most enterprises focus on edge computing as it minimizes delays, which is one of the major factors for any real-time application to perform efficiently.
- **AI growth favours cloud growth**: New AI-oriented services and technology are helping the major cloud providers to ride a wave. New capabilities lead to increased demand, which leads to increased revenues, which then enables more investments

Source:
https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/

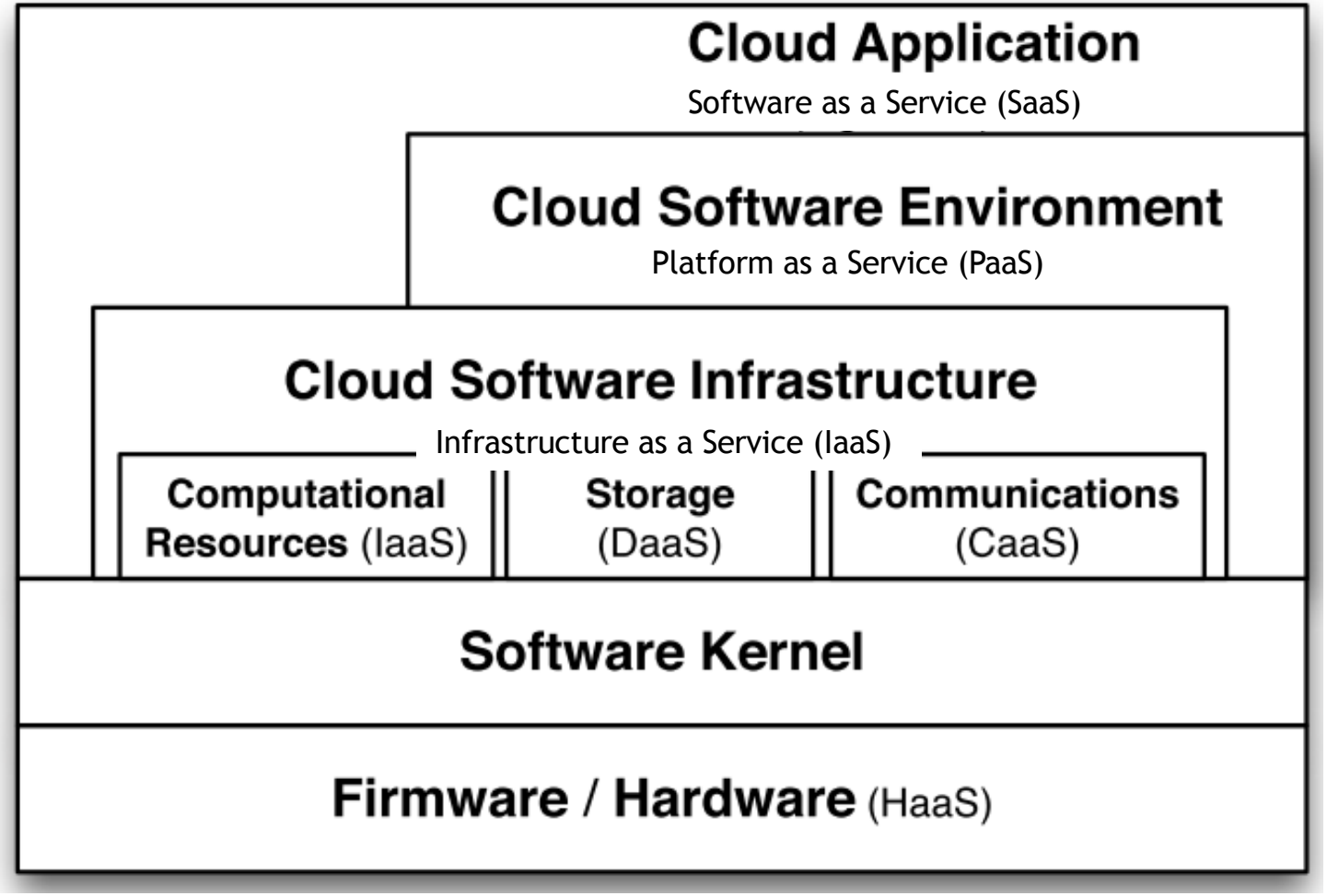
A variety of ‘as-a-Service’ terms to describe services offered in Clouds - XaaS

AaaS	-	Architecture as a Service
BaaS	-	Business as a Service
CaaS	-	Communication as a Service
CRMaaS	-	CRM as a Service
DaaS	-	Data as a Service
DBaaS	-	Database as a Service
EaaS	-	Ethernet as a Service
FaaS	-	Function as a Service
GaaS	-	Globalization or Governance as a Service
HaaS	-	Hardware as a Service
IaaS	-	Infrastructure as a Service
IDaaS	-	Identity as a Service
ITaaS	-	IT as a Service
LaaS	-	Lending as a Service
MaaS	-	Mashups as a Service
OaaS	-	Organization or Operations as a Service
SaaS	-	Software as a Service
StaaS	-	Storage as a Service
PaaS	-	Platform as a Service
TaaS	-	Technology or Testing as a Service
VaaS	-	Voice as a Service

Three main services provided by Cloud …

"Toward a Unified Ontology of Cloud Computing"

[*L. Youseff, M. Butrico, and D. Da Silva*]



Cloud Application Layer

Cloud Application Layer

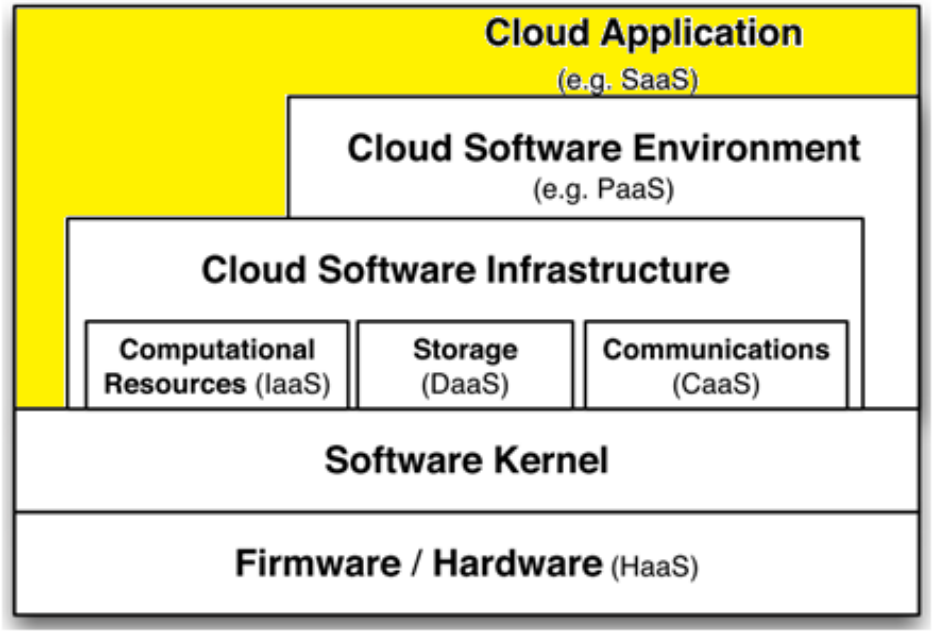
- **SaaS**

Users access the services provided by this layer through web-portals and are *sometimes* required to pay fees to use them

Cloud applications can be developed on the cloud software environments or infrastructure components

Example:

- GMail
- Google Docs and related apps (online office)
- SalesForce.com (CRMaaS)



Cloud Software Environment Layer

Cloud Software Environment Layer

- **PaaS**

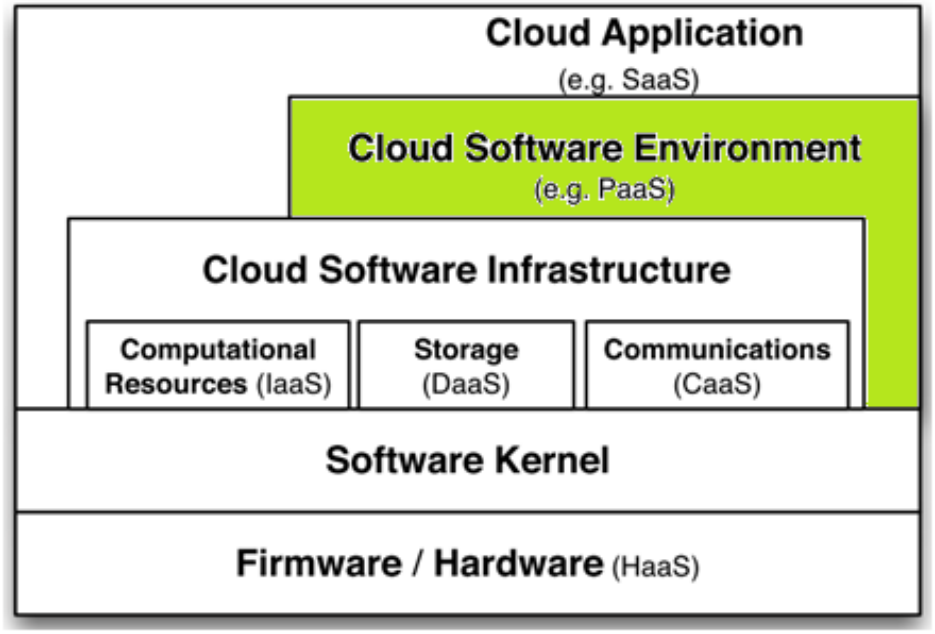
Users are *application developers*

Providers supply developers with a *programming-language-level* **environment** with a well-defined **API**

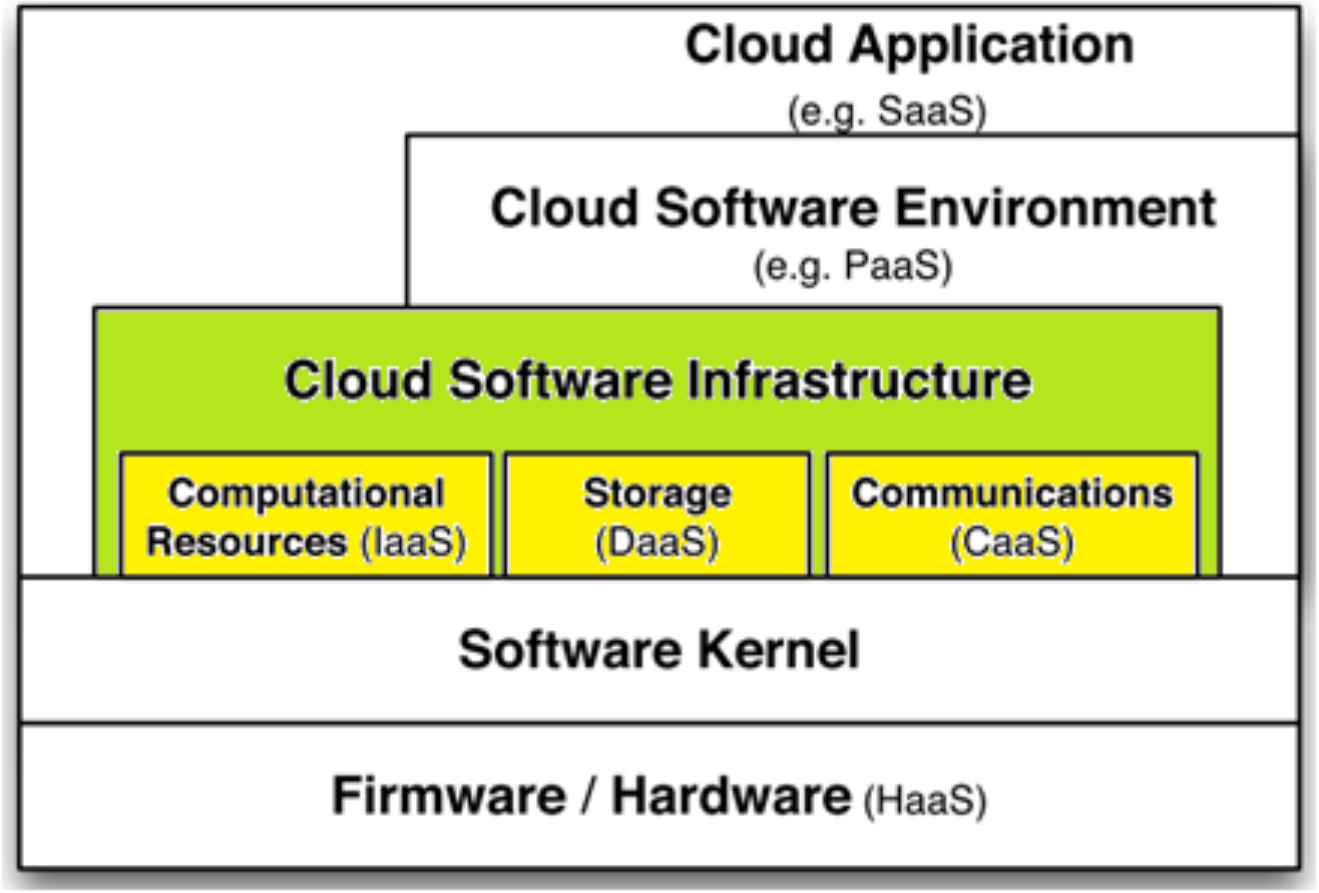
- Facilitate interaction between environment and apps
- Accelerate the deployment
- Support scalability

Examples in Deep Learning:

- Amazon SageMaker
- Microsoft Azure Machine Learning
- Google AI: TensorFlow



Cloud Software Infrastructure Layer



Cloud Software Infrastructure Layer

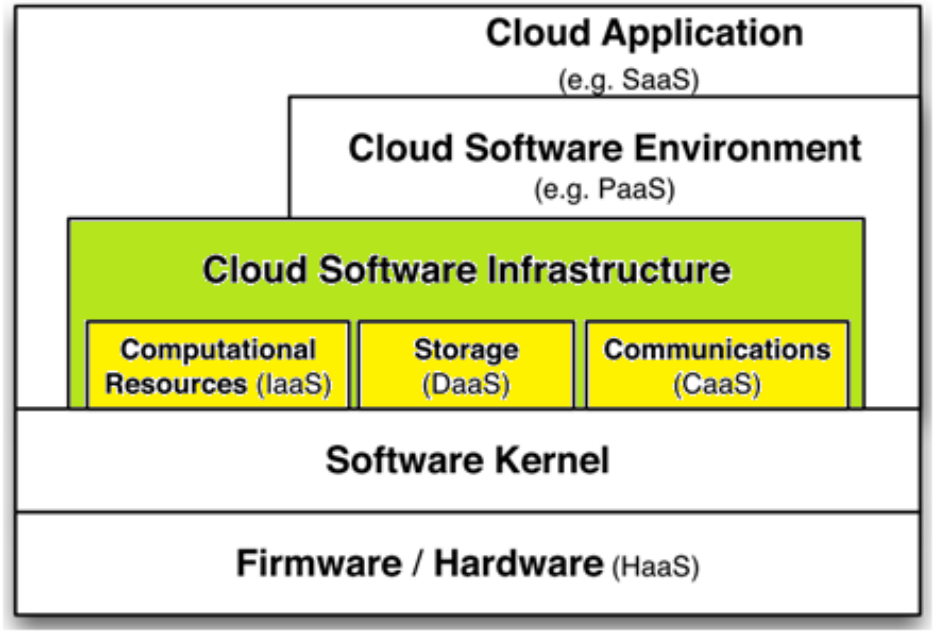
Cloud Software Infrastructure Layer

- **IaaS**: computational
- **DaaS**: storage
- **CaaS**: communications

Provides resources to the higher-level layers (i.e., Application and Software Environment)

Note that Cloud Apps and Cloud SW might *bypass* Cloud SW Infrastructure

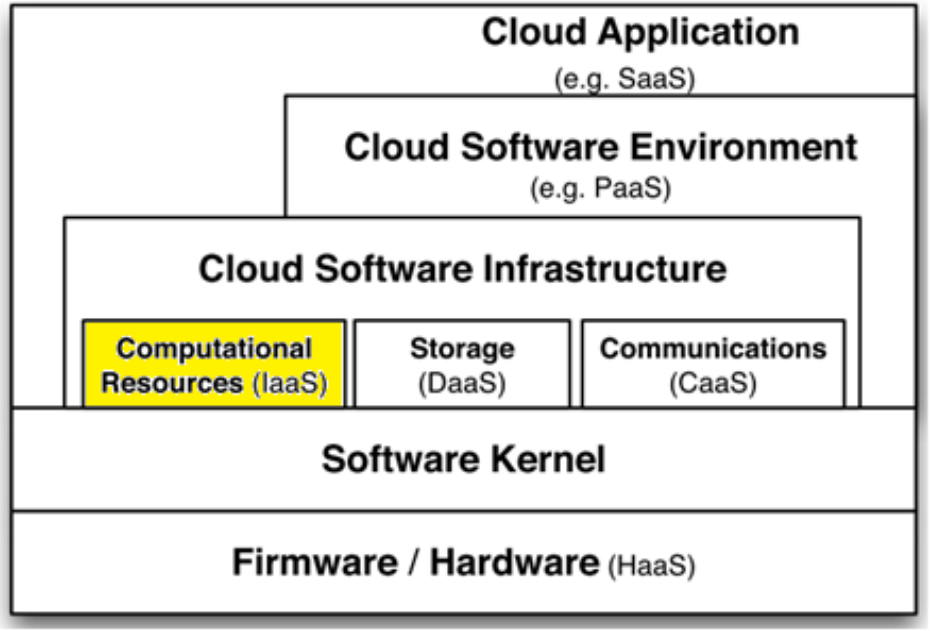
- However, this would:
  - reduce simplicity
  - increase development efforts



Infrastructure as a Service (IaaS)

Virtual Machines (VM) vs dedicated hardware

- VM's benefits
  - Flexibility
  - Super-user (root) access to VM for fine granularity settings and customization of installed sw
- VM's issues
  - Performance interference
  - Inability to provide strong guarantees about SLAs



Infrastructure as a Service (IaaS): examples

- **Commercial solutions**
  - **Amazon Elastic Cloud (EC2)**
    - Full virtualization
    - Based on Xen
  - **Windows Azure**
    - Not just windows-based: it allows also to start VMs for other OSs
  - **Google Compute Engine**
    - Same infrastructure as Google
  - **Rackspace Open Cloud**
  - **IBM SmartCloud Enterprise**
  - **HP Enterprise Converged Infrastructure**
- **Open-source projects**
  - **Eucalyptus** Systems
  - **Apache CloudStack**
  - **OpenStack**
    - The project aims to deliver solutions for all types of clouds (private or public) by being simple to implement, massively scalable, and feature rich

Data as a Service (DaaS)

Allows users to

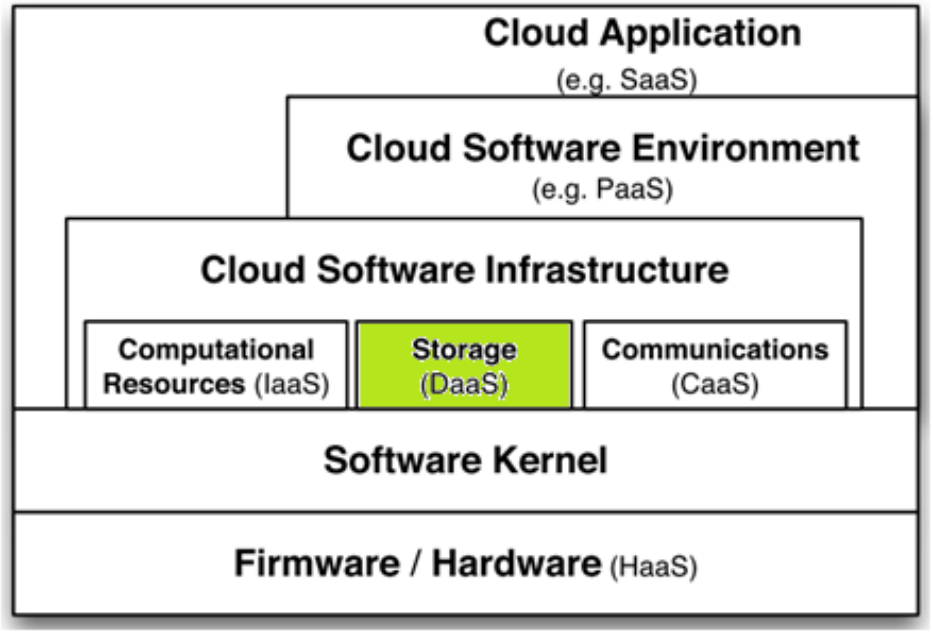
- store their data at remote disks
- access data anytime from any place

Facilitates cloud applications to scale beyond their limited servers requirements:

- High dependability: availability, reliability, performance (scalability)
- Replication
- Data consistency

DropBox, iCloud, GoogleDrive are examples of DaaS

CEPH is an open source solution

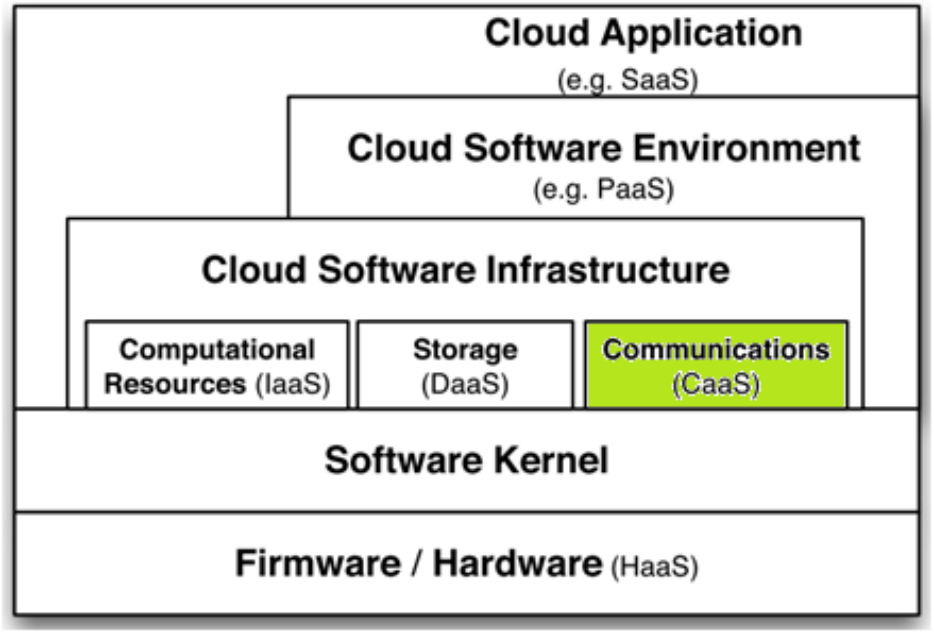


Communications as a Service (CaaS)

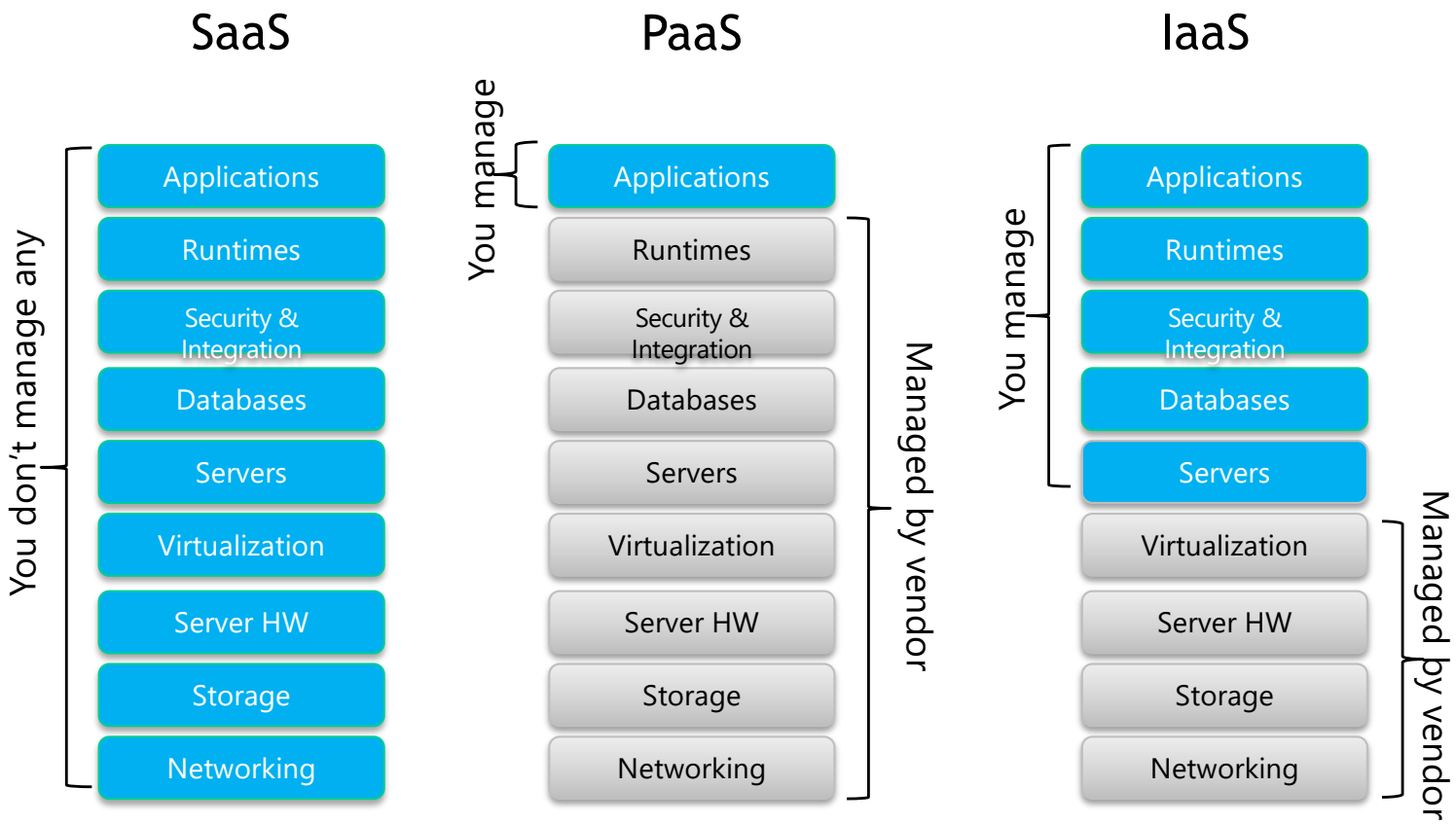
Communications becomes a vital component in guaranteeing QoS

- Communication capability: service oriented, configurable, schedulable, predictable, and reliable
- Network security, dynamic provisioning of virtual overlays for traffic isolation or dedicated bandwidth, guaranteed message delay, communication encryption, and network monitoring

Types of CaaS include Voice over Internet Protocol (VoIP) or internet telephone solutions, and video conferencing services

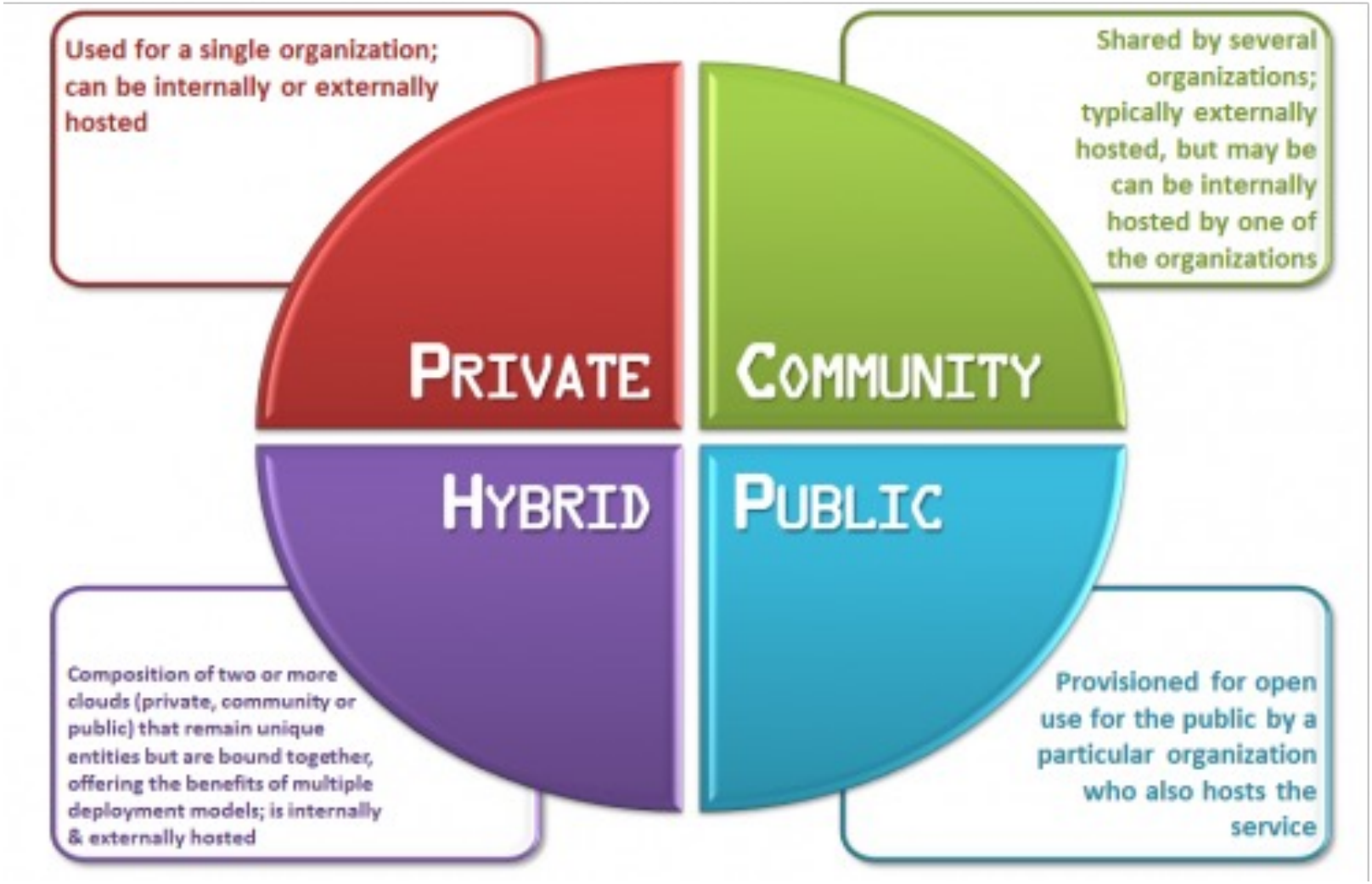


SaaS, PaaS, IaaS summary



Courtesy of Microsoft

Types of Clouds



Public Clouds

Large scale infrastructure available on a rental basis

- The definition of Cloud we gave so far

Fully customer self-service

- Service Level Agreements (SLAs) are advertized
- Requests are accepted and resources granted via web services
- Customers access resources remotely via the Internet

Accountability is e-commerce based

- Web-based transaction
- “Pay-as-you-go” and flat-rate subscription
- Customer service, refunds, etc.

Private Clouds

Internally managed data centers

The organization sets up a **virtualization** environment on its **own servers**

- in its data center
- in the data center of a managed service provider

Key benefits

- you have **total control over every aspect** of the infrastructure
- you gain advantages of virtualization

Issues

- it lacks the freedom from
  - capital investment
  - flexibility (“almost infinite” grow of cloud computing)

Useful for companies that have significant existing IT investments

Community Clouds

A single cloud managed by several federated organizations

- Combining together several organizations allows economy of scale
- Resources can be shared and used by one organization, while the others are not using them

Technically similar to private cloud:

- They share the same software and the same issues
- A more complex accounting system is however required

Hosted locally or externally:

- Typically community clouds shares infrastructures of the participants
- However they can be hosted by a separate specific organization, or only by a small subset of the partners

Hybrid Clouds

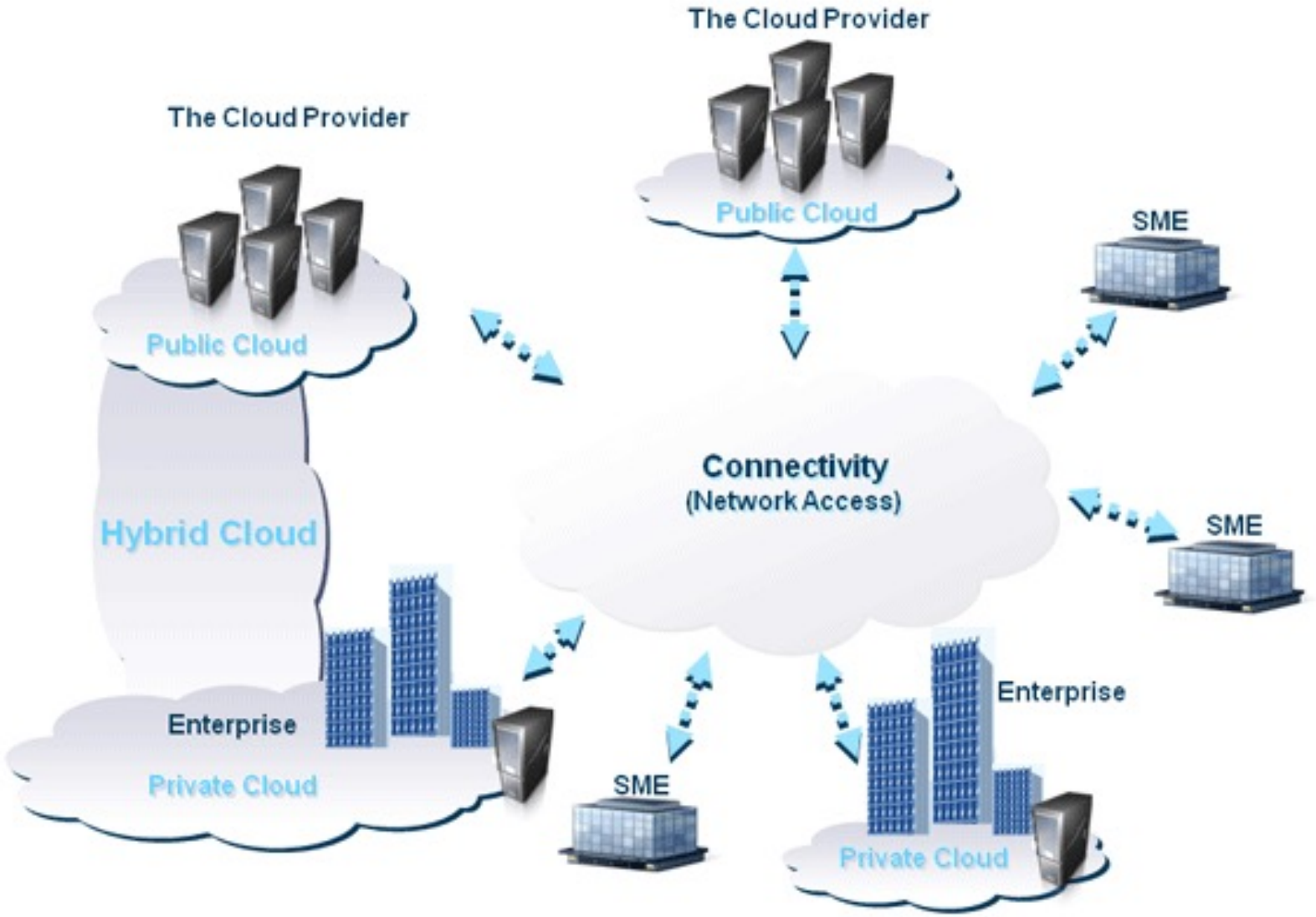
Hybrid clouds are the combination of any of the previous types.

- Usually are companies that holds their private cloud, but that they can be subject to unpredictable peaks of load
- In this case, the company rents resources from other types of cloud

Common interfaces

- To simplify the deployment process, the way in which VMs are started, terminated, address is given and storage is accessed, must be as similar as possible
- Many standards are being developed in this directions, but none is globally accepted yet
- Currently, the Amazon EC2 model is the one with more compliant infrastructures

Types of Cloud







Cloud Computing - The commercial case: Amazon EC2

Elastic Compute Cloud

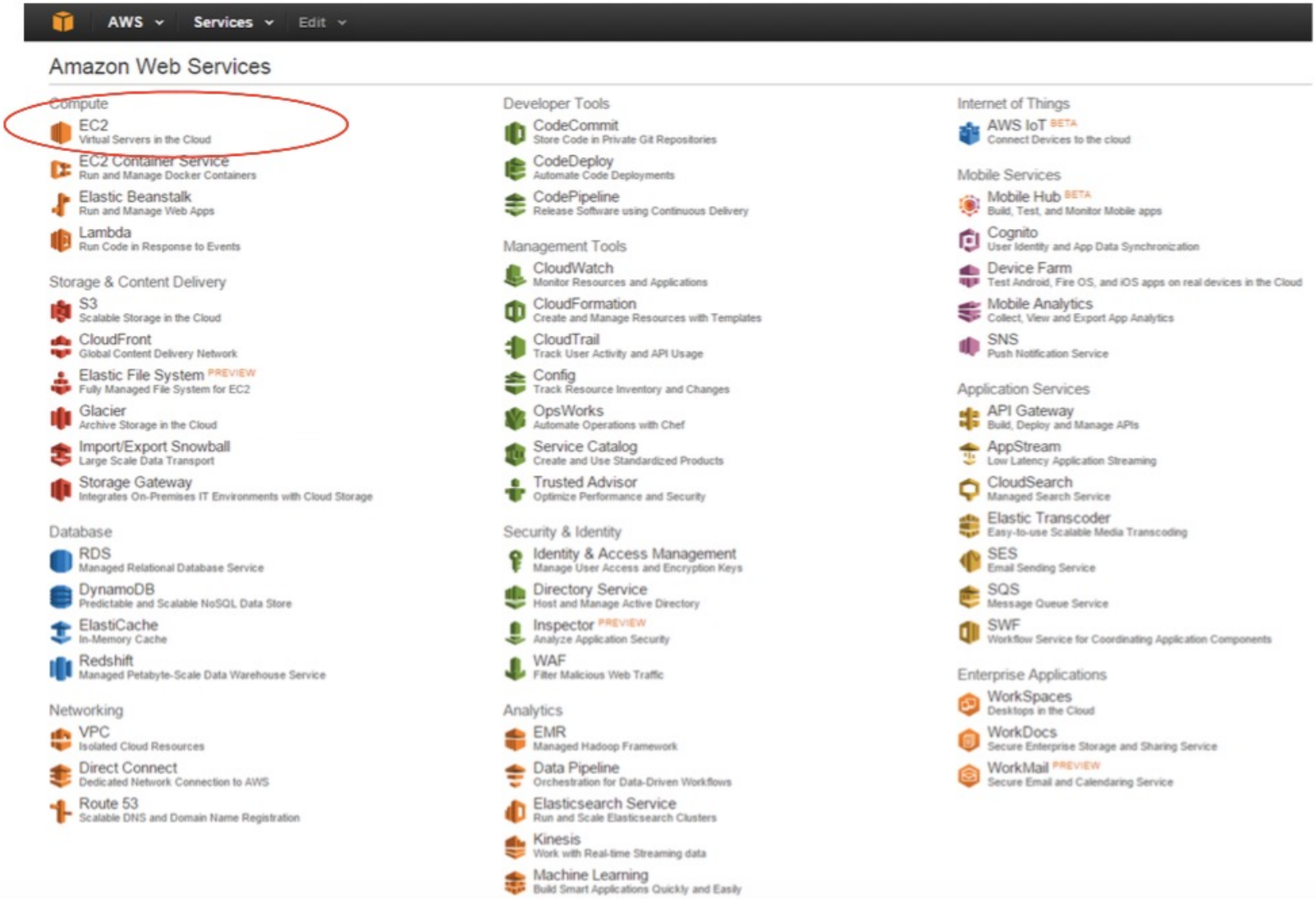
Rent virtual machine instances to run your software. Monitor and increase / decrease the number of VMs as demand changes

How to use:

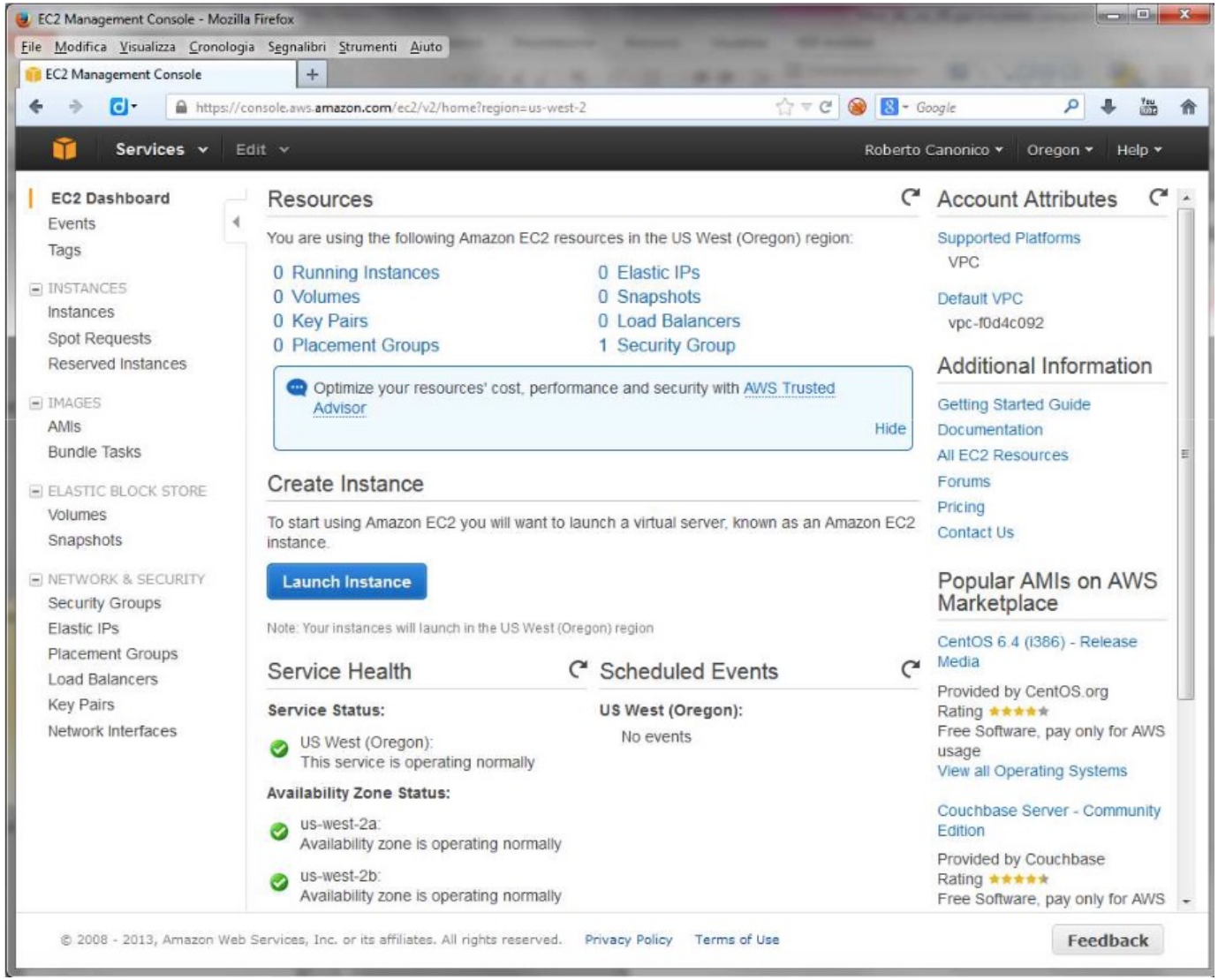
- Create an Amazon Machine Image (AMI): applications, libraries, data and associated settings
- Upload AMI to Amazon S3 (simple storage service)
- Use Amazon EC2 web service to configure security and network access
- Choose OS, start AMI instances
- Monitor & control via web interface or APIs

Tutorial:
https://aws.amazon.com/getting-started/launch-a-virtual-machine-B-0/

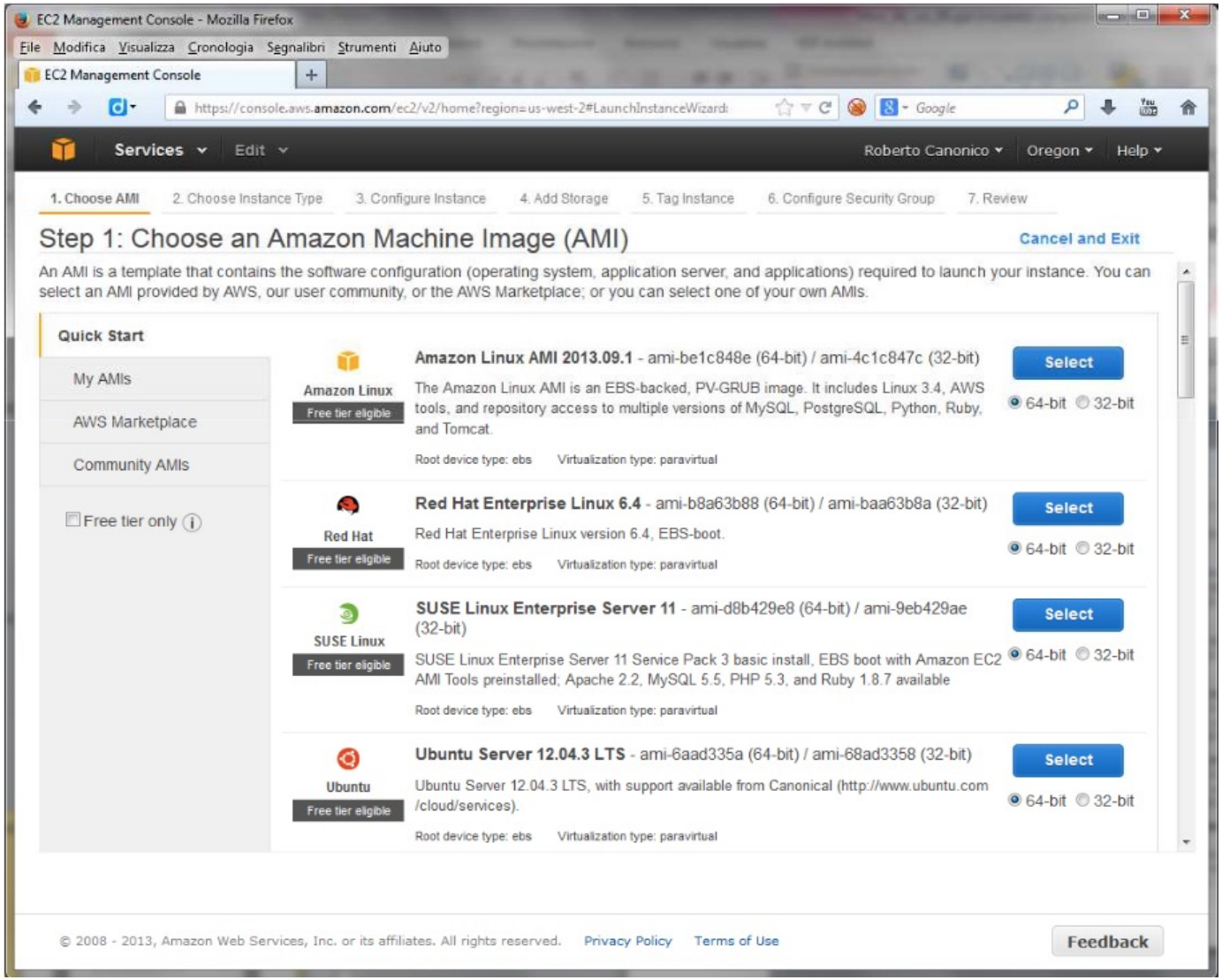
EC2 is an Amazon Web Service



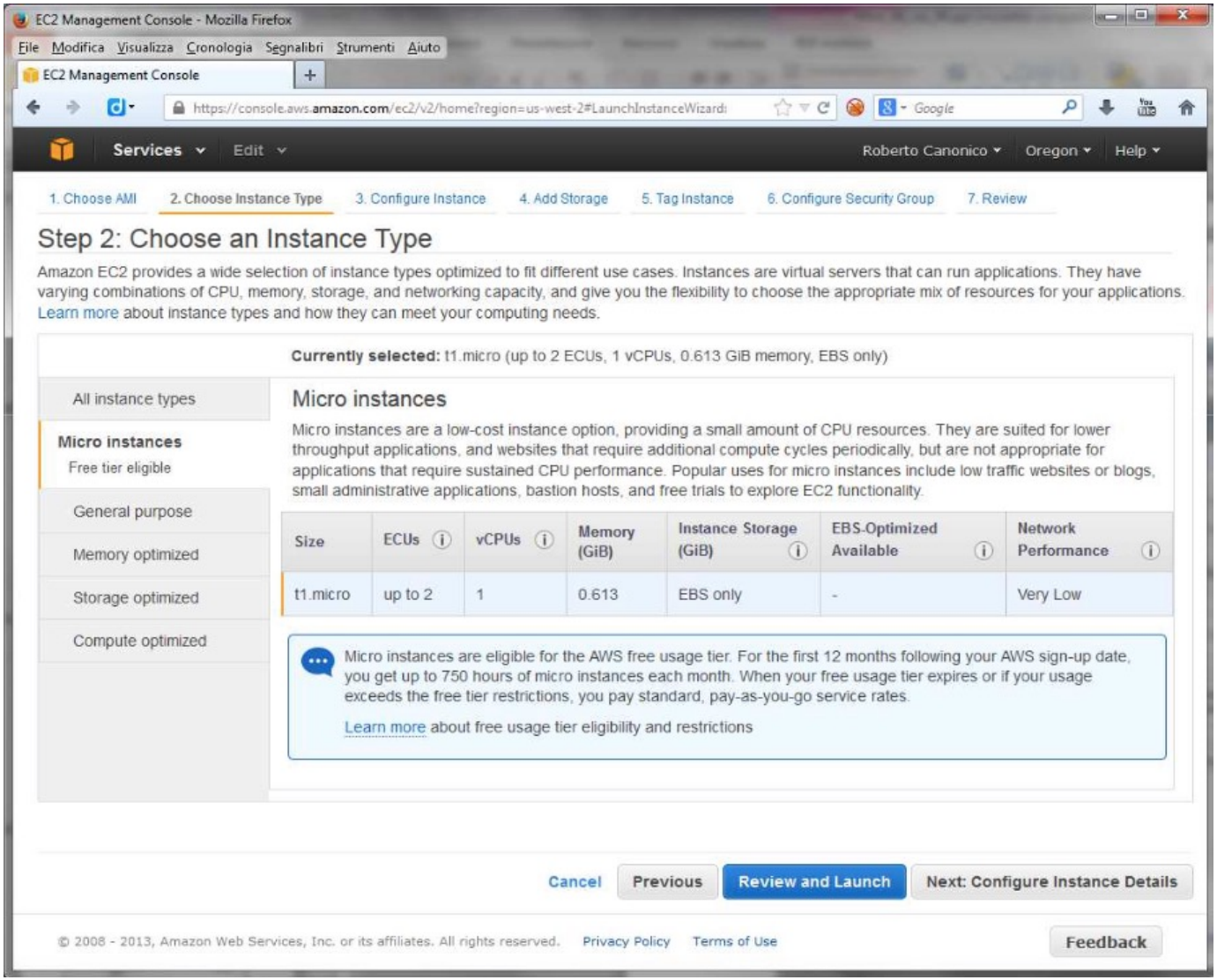
Amazon EC2: the console



Amazon EC2: selecting the AMI



Amazon EC2: creating an instance



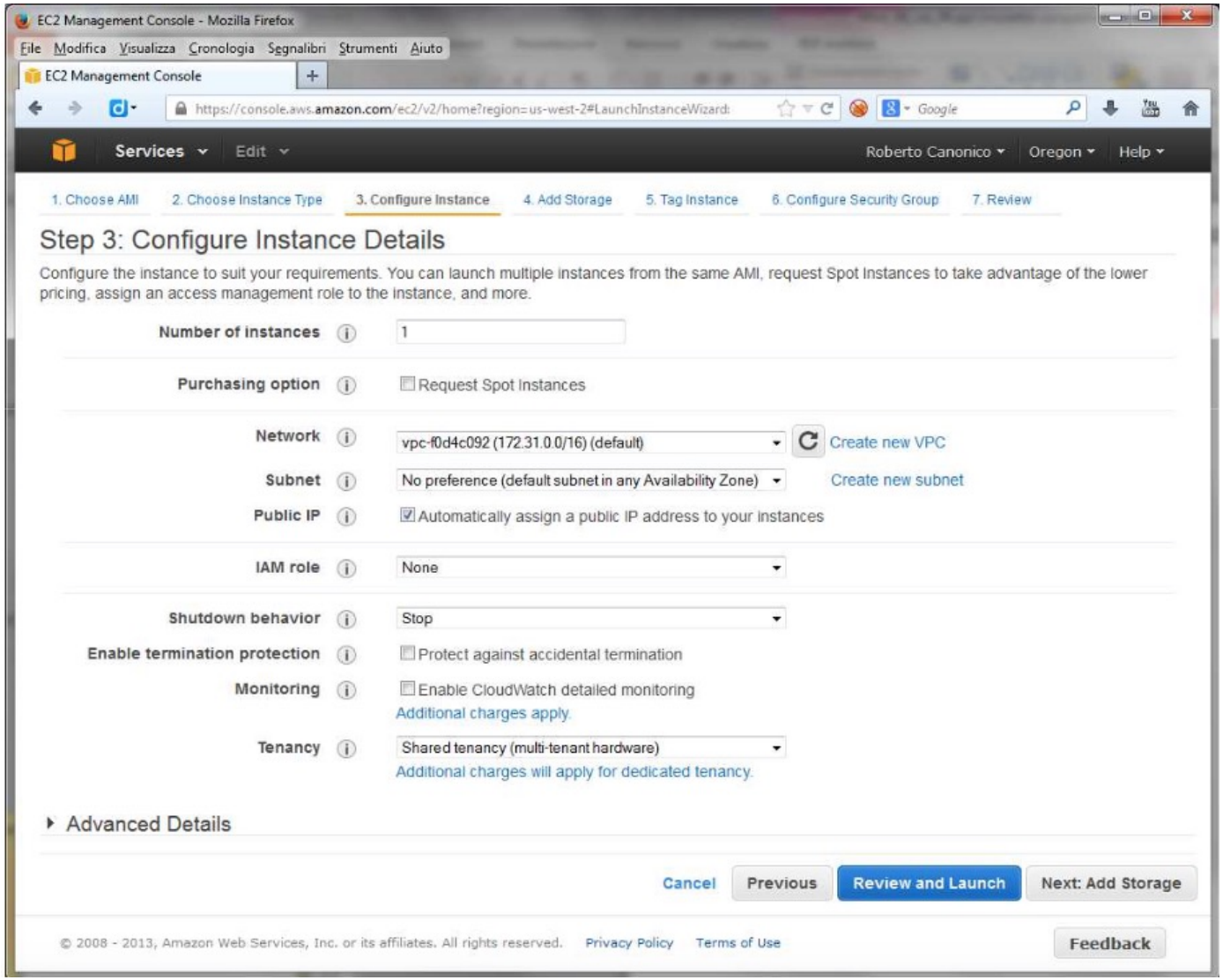
Amazon instance types

Instances are divided into *types*, that corresponds to different performance characteristics, and different pricing.

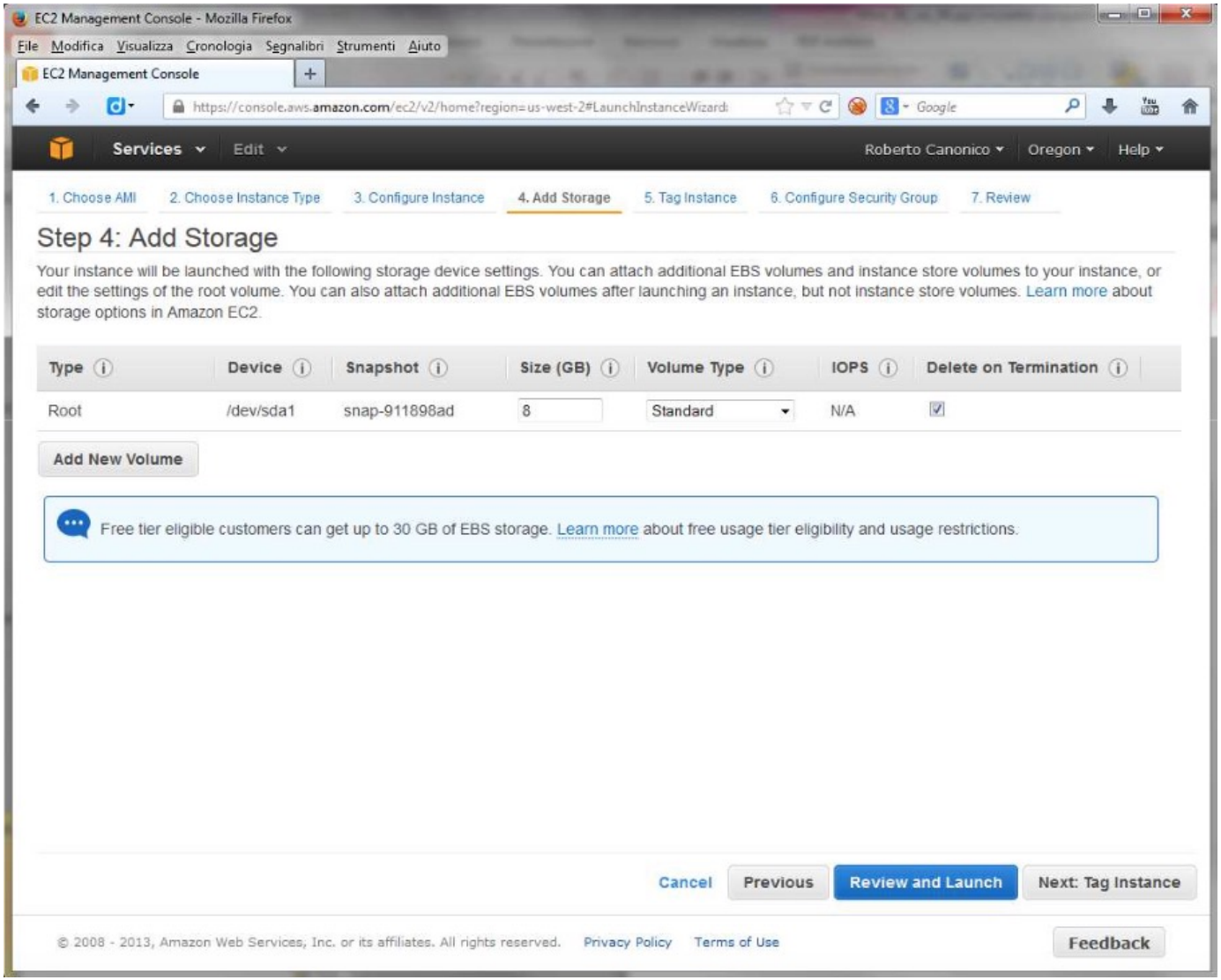
Speed of the instances is measured in vCPU: an Hyperthread of an Intel Xeon core.

	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage
General Purpose - Current Generation					
t2.nano	1	Variable	0.5	EBS Only	$0.0059 per Hour
t2.micro	1	Variable	1	EBS Only	$0.012 per Hour
t2.small	1	Variable	2	EBS Only	$0.023 per Hour
t2.medium	2	Variable	4	EBS Only	$0.047 per Hour
t2.large	2	Variable	8	EBS Only	$0.094 per Hour
t2.xlarge	4	Variable	16	EBS Only	$0.188 per Hour
t2.2xlarge	8	Variable	32	EBS Only	$0.376 per Hour
m4.large	2	6.5	8	EBS Only	$0.108 per Hour
m4.xlarge	4	13	16	EBS Only	$0.215 per Hour
m4.2xlarge	8	26	32	EBS Only	$0.431 per Hour
m4.4xlarge	16	53.5	64	EBS Only	$0.862 per Hour
m4.10xlarge	40	124.5	160	EBS Only	$2.155 per Hour
m4.16xlarge	64	188	256	EBS Only	$3.447 per Hour

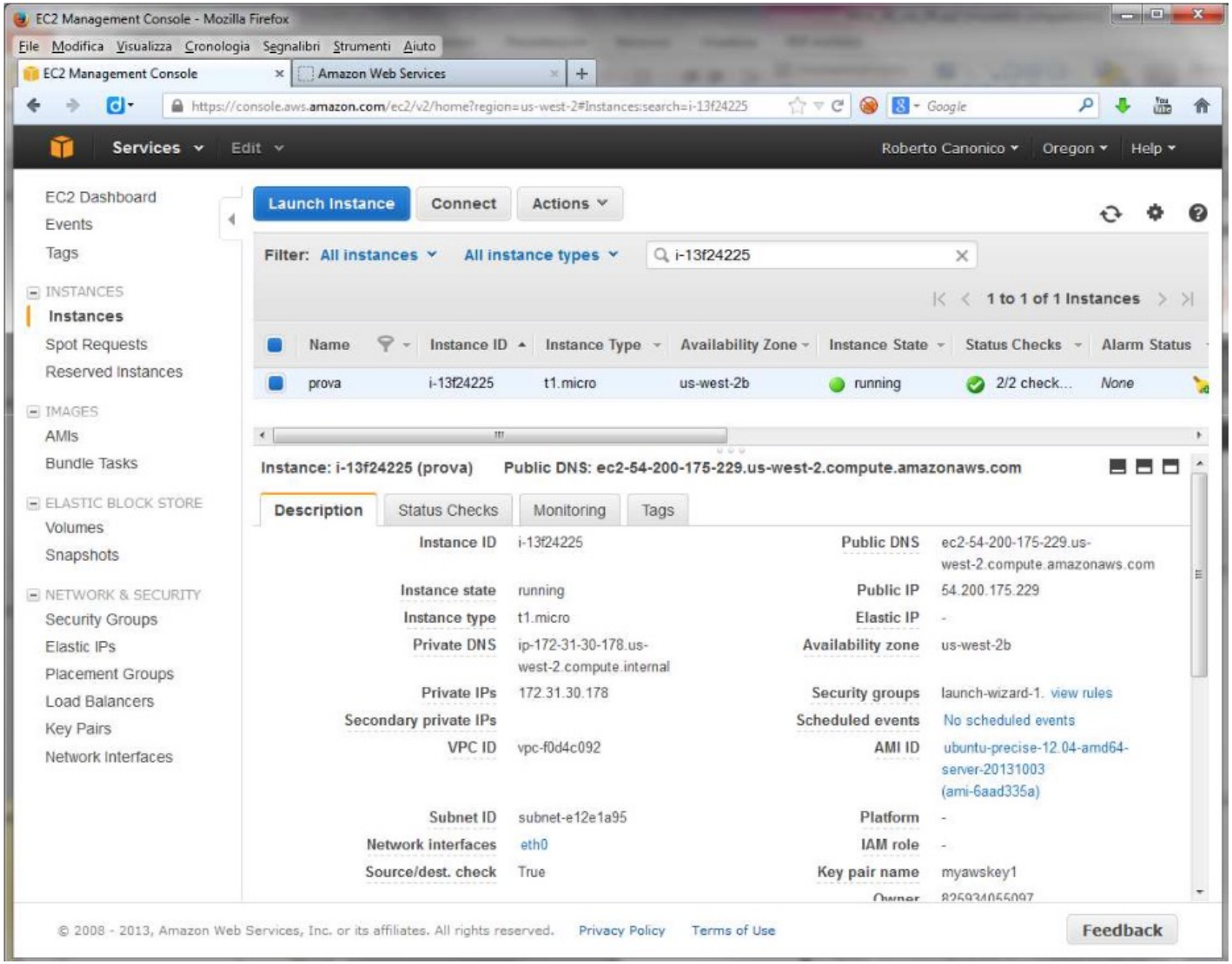
Amazon EC2: configuring an instance (2)



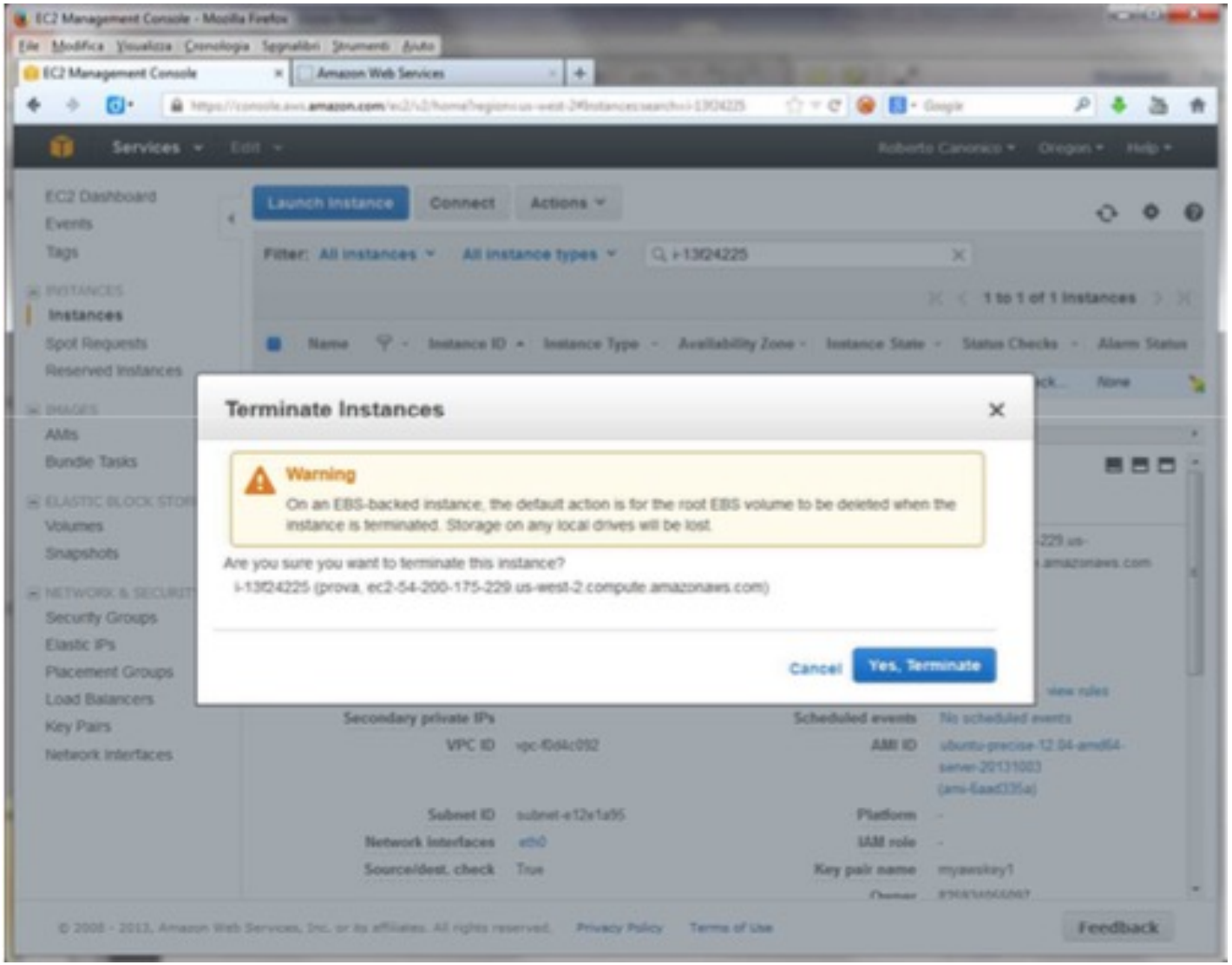
Amazon EC2: adding a storage



Amazon EC2: launching the instance(s)



Amazon EC2: instance termination



POLITECNICO DI MILANO





From Cloud to Edge and Fog Computing

Advantages of Cloud Computing

- Lower IT costs
- Improved performance
- Instant software updates
- “Unlimited” storage capacity
- Increased data reliability
- Universal document access
- Device Independence

What about disadvantages?

Disdvantages of Cloud Computing

- Requires a constant Internet connection
- Does not work well with low-speed connections
- Can be slow
- Features might be limited
- Stored data might not be secure
- Lock-in

Fog/Edge Computing

- When it comes to storage and computation of large scales of data, Cloud Computing is the de-facto solution
- With the massive growth in intelligent and mobile devices coupled with technologies like Internet of Things (IoT), V2X Communications, Augmented Reality (AR), the focus has shifted towards
  - gaining real-time responses
  - support for context-awareness
  - mobility
- Due to the delays induced on the WAN and location agnostic provisioning of resources on the cloud, there is a need to bring the features of the cloud closer to the consumer devices
- Nowadays, computing/storage capacity are available at the data sources

Fog/Edge Computing