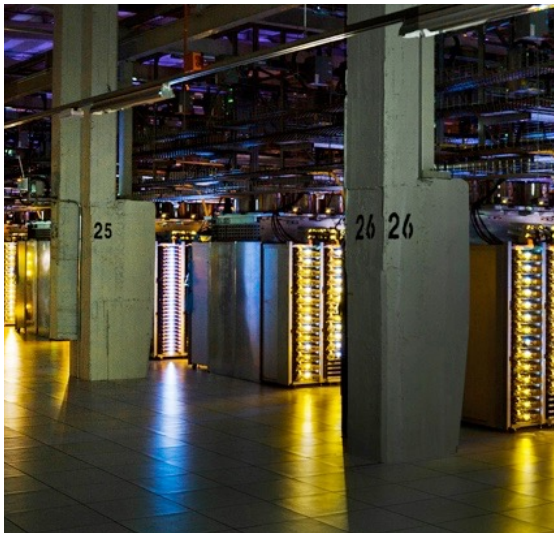
Cloud Computing — Introduction to Amazon Web Services



Academic year 2014/15

Amazon Web Services
Introduction

- Amazon Web Services (AWS) is a collection of remote infrastructure services mainly in the *Infrastructure as a Service* (IaaS) category, with some services in the *Platform as a Service* (PaaS) category.

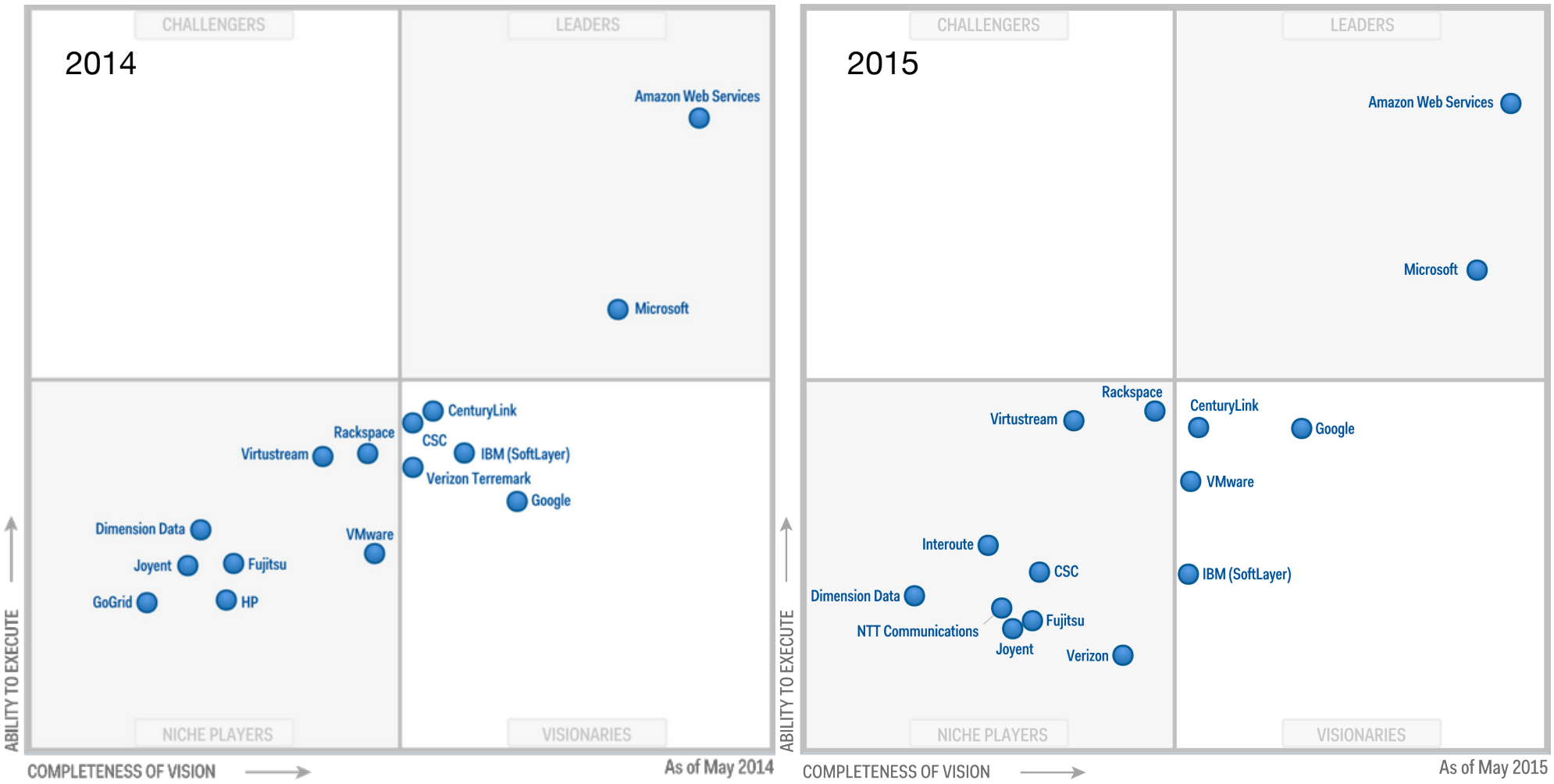
- Introduced in 2006/2007, it is considered the first real cloud computing offering.

- In IaaS the main services offered by AWS are
  - **Compute** — for example *Elastic Compute Cloud*
  - **Storage** — for example *Simple Storage Service*
  - **Database** — for example *Relational Database Service*
  - **Networking** — for example *Virtual Private Cloud*
  - ...

- The services are targeted towards *operations engineers* and *developers*.

Amazon Web Services

Gartner Magic Quadrant for Cloud Infrastructure as a Service



Source: Gartner

Amazon Web Services

Introduction — What was Amazon's motivation in creating AWS?

- Typical weekly traffic to Amazon's e-commerce web site in 2007

Amazon Web Services

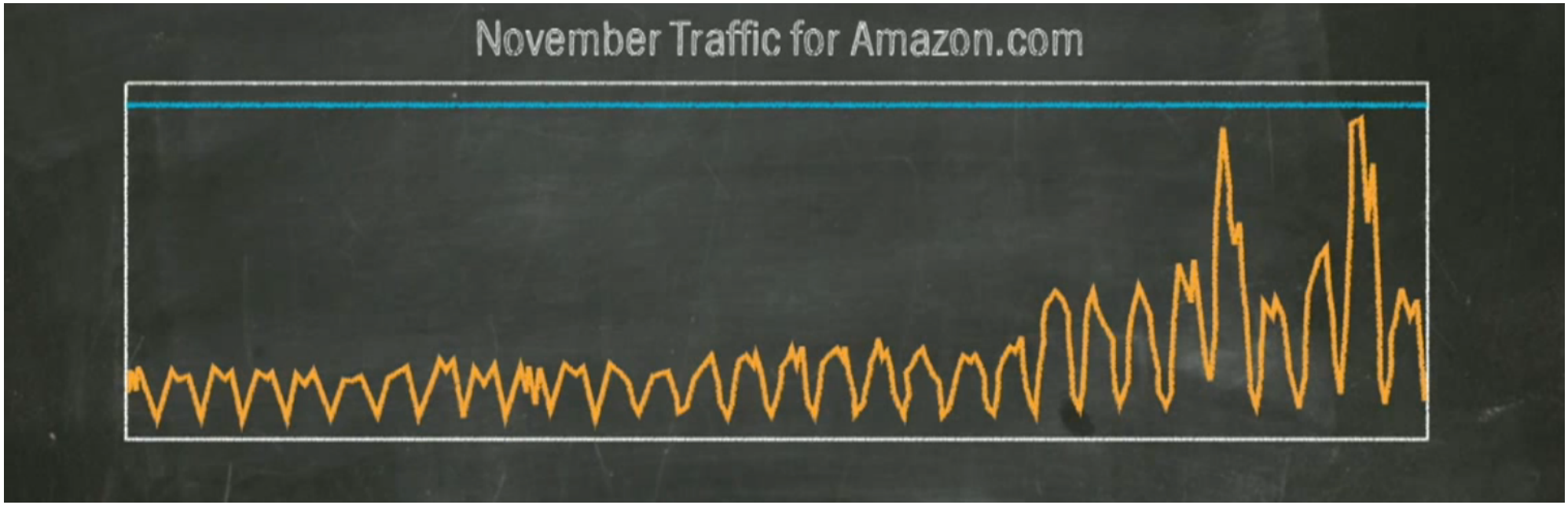
Introduction — What was Amazon's motivation in creating AWS?

- Typical weekly traffic to Amazon's e-commerce web site in 2007

Amazon Web Services

Introduction — What was Amazon's motivation in creating AWS?

- Traffic in the month of November 2007



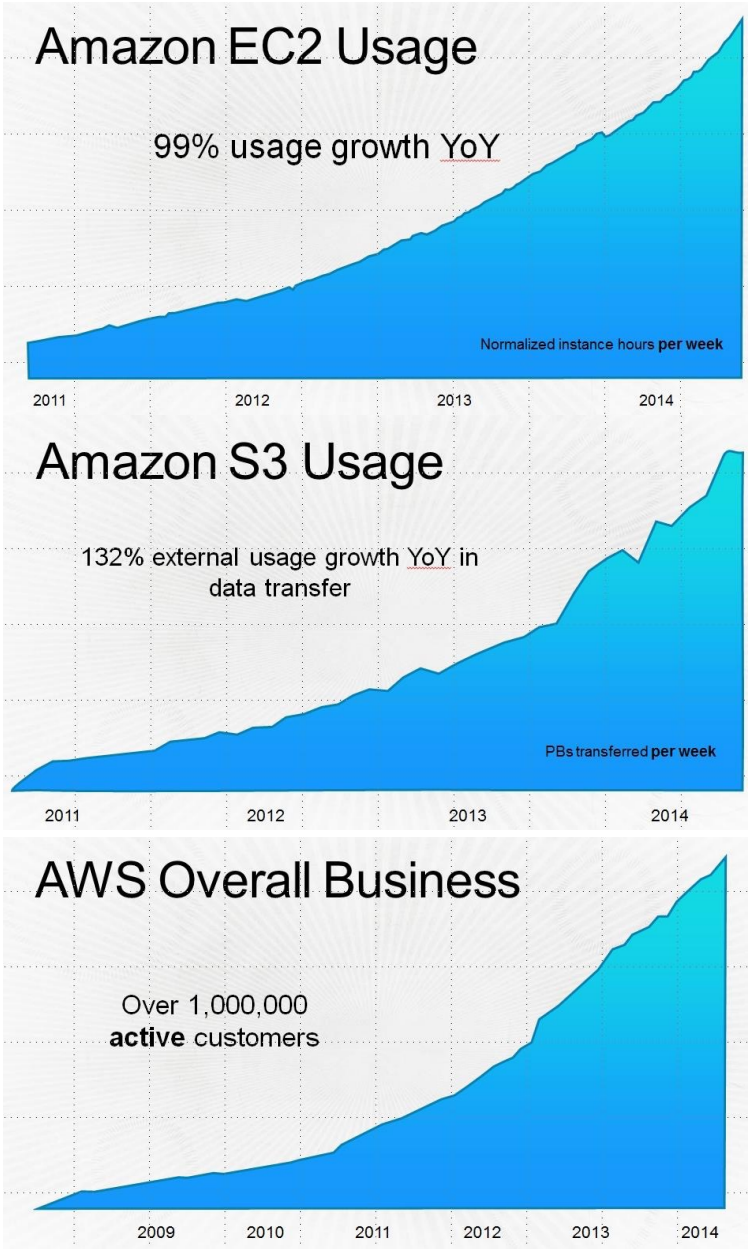
Amazon Web Services

Introduction — What was Amazon's motivation in creating AWS?

- Traffic in the month of November 2007

Amazon Web Services
AWS today

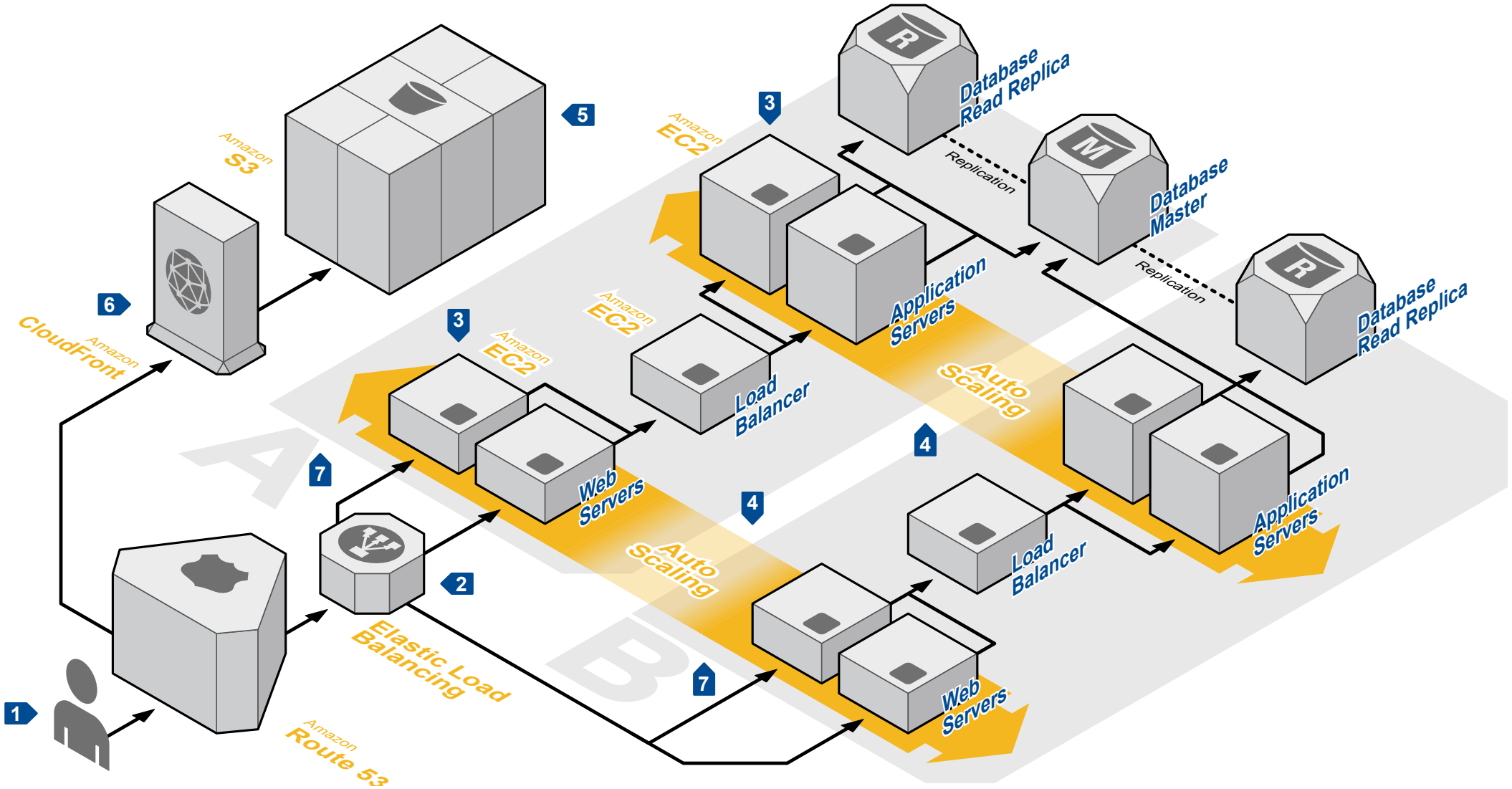
- Very big
  - $5B annual revenue
  - “5 times the cloud capacity in use than the aggregate total of the other 14 providers” *Gartner*
- Growing very fast
  - “Every day, AWS adds enough new server capacity to support all of Amazon’s global infrastructure when it was a $7B annual revenue enterprise” *Amazon*



Source: Amazon

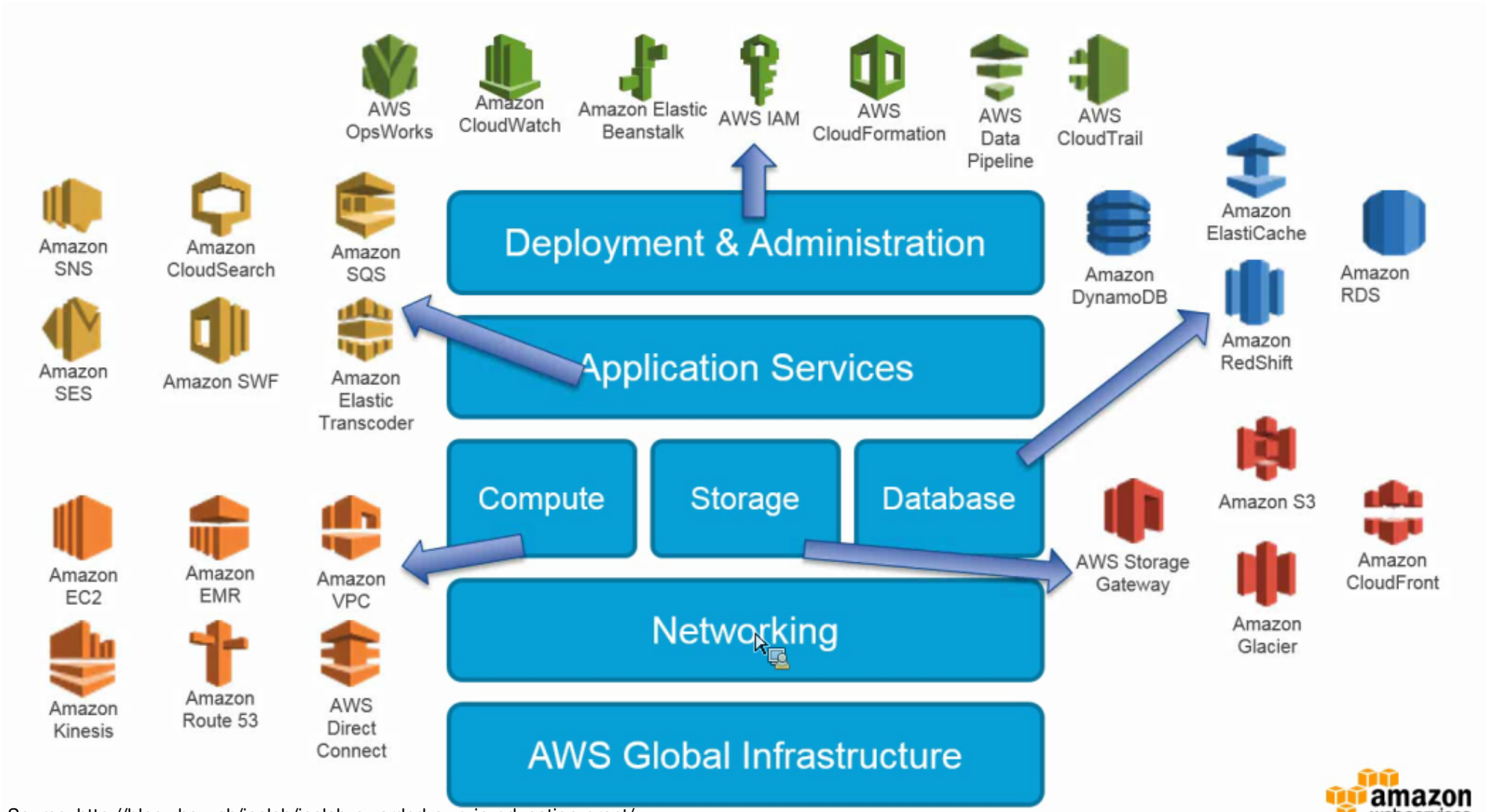
Amazon Web Services

Example: Hosting of a web application



Amazon Web Services

Service stack

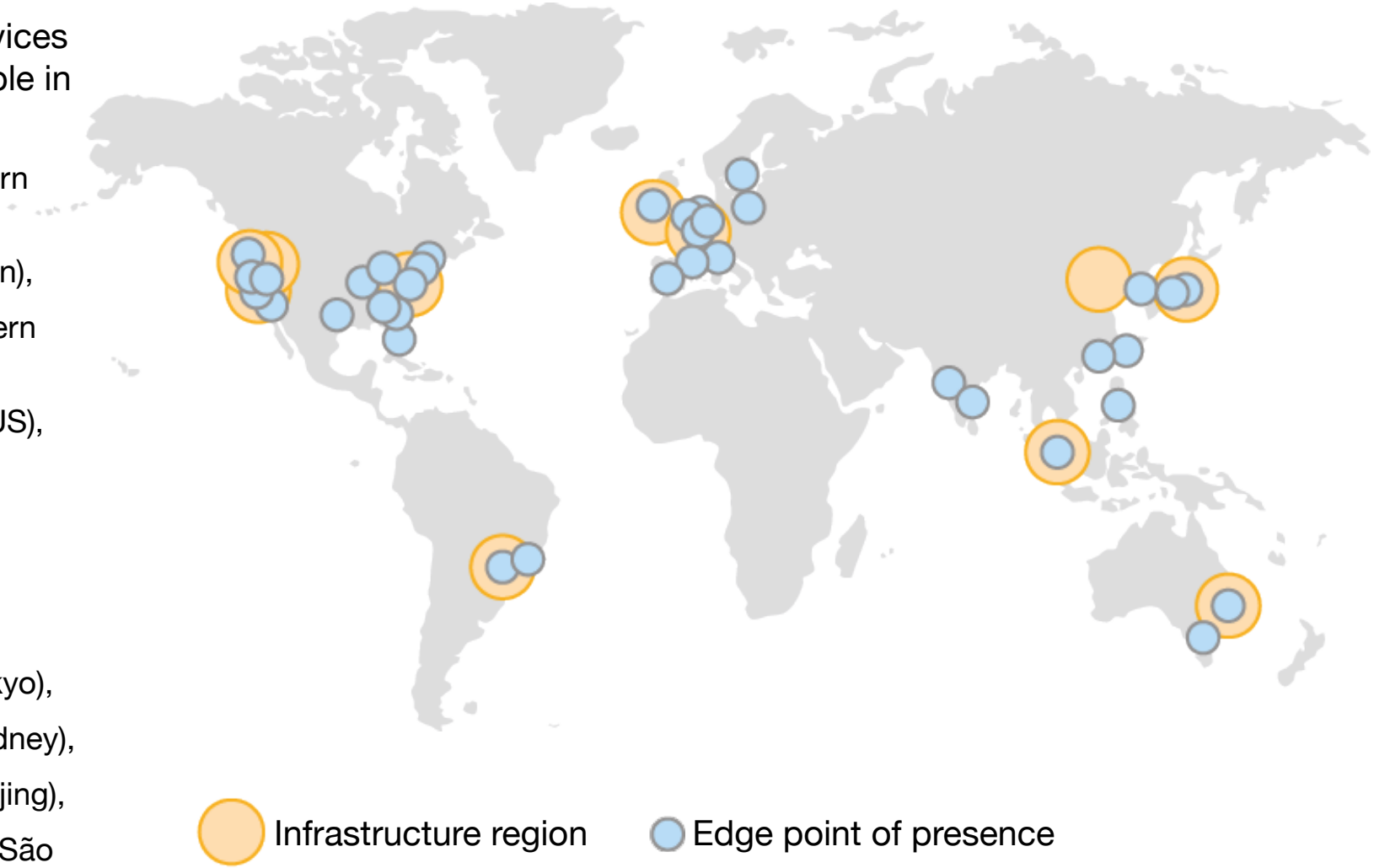


Source: http://blog.zhaw.ch/icclab/icclab-awarded-aws-in-education-grant/

Amazon Web Services

Global datacenters

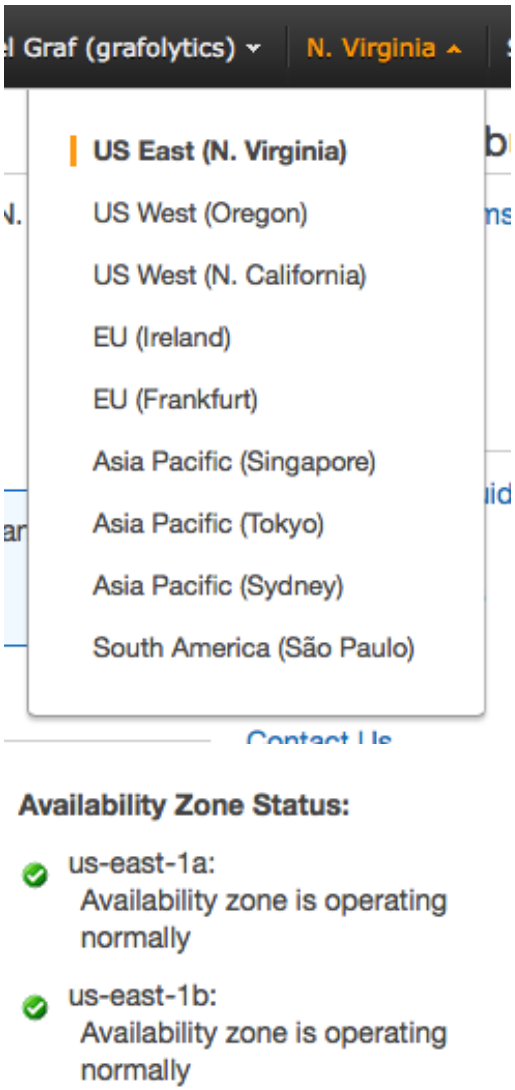
- Amazon Web Services is currently available in eleven regions:
  - US East (Northern Virginia),
  - US West (Oregon),
  - US West (Northern California),
  - US GovCloud (US),
  - EU (Ireland),
  - EU (Frankfurt),
  - Asia Pacific (Singapore),
  - Asia Pacific (Tokyo),
  - Asia Pacific (Sydney),
  - Asia Pacific (Beijing),
  - South America (São Paulo)



Amazon Web Services

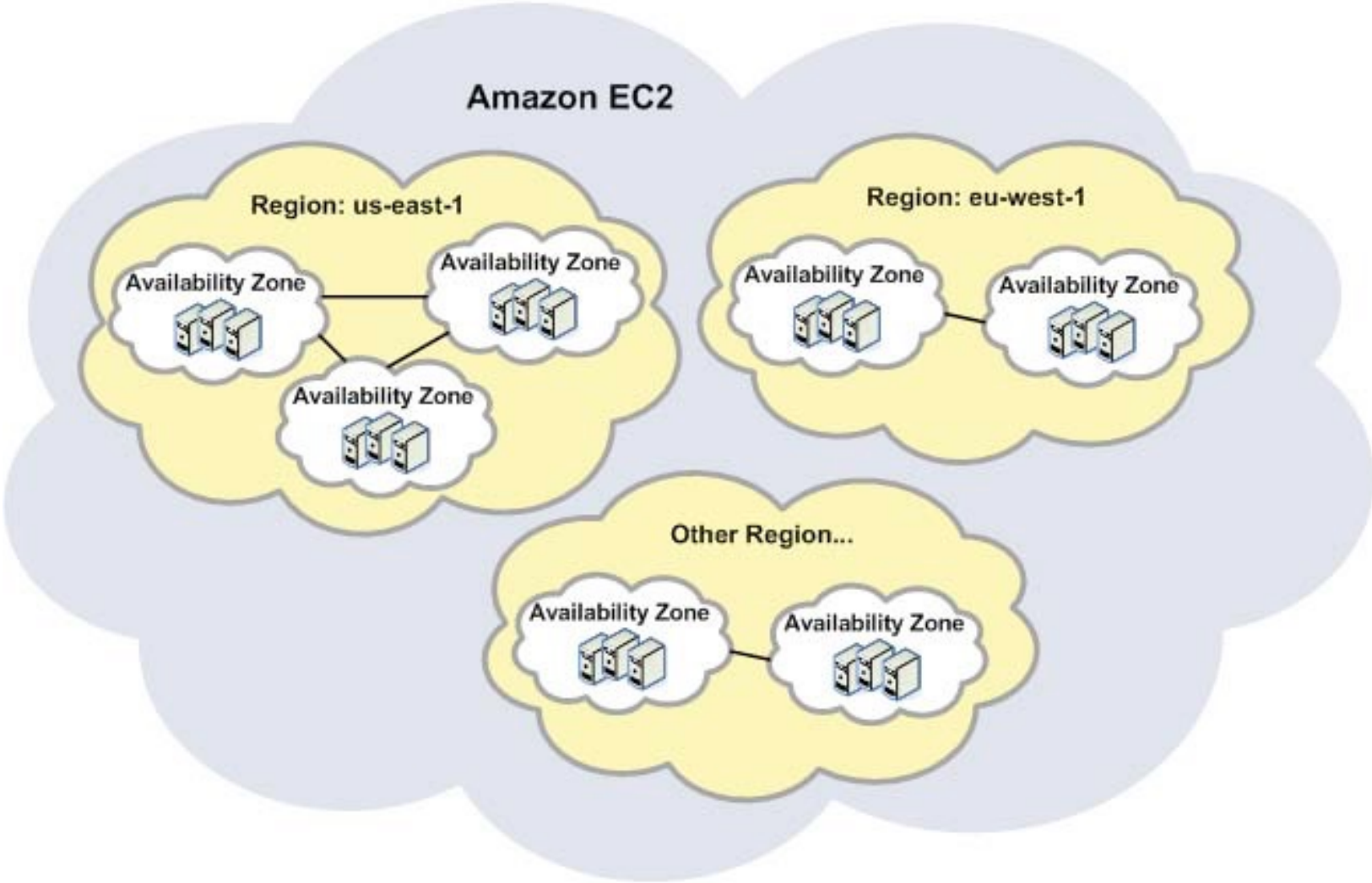
Regions and *Availability Zones*

- When allocating a cloud resource (e.g., a virtual machine), the AWS customer can choose the *region*, and within the region, the *availability zone*.
- **Regions** are distributed globally and enable a developer to place his application and/or data in a particular country / region
  - to be closer to his customers
  - so that the data resides in a particular jurisdiction to be compliant with regulations (e.g., data privacy laws)
- **Availability Zones** are separate datacenters inside a Region
  - Each datacenter has its own independent infrastructure for power, cooling, … When disaster strikes its effects may be limited to a single Availability Zone.
  - Availability Zones in a region are connected with low-latency network links
  - An AWS customer can distribute redundant copies of his application over several Availability Zones to protect it from a single failing Availability Zone.



Amazon Web Services

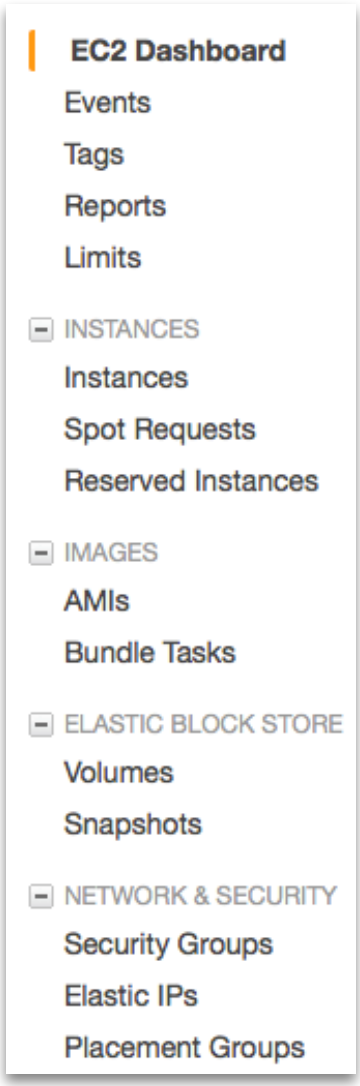
Regions and *Availability Zones*



Amazon Web Services

Elastic Compute Cloud (EC2)

- Amazon *Elastic Compute Cloud* (EC2) is a web service that provides resizable compute capacity in the cloud.
- Publicly available since 2006
- Is considered the first real cloud computing product.
- Developers can rent virtual machines (called *EC2 instances*) by the hour
- Many *instance types* are available



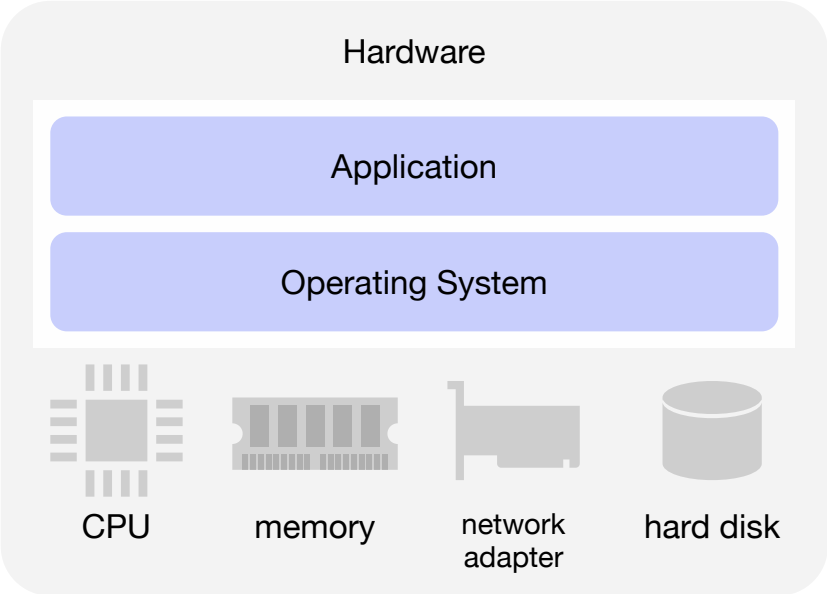
Virtualization
Introduction

- Virtualization in Computer Science: create, for a program or a process, the appearance of a resource (virtual resource) that does not exist in this form physically.
  - Virtual memory managed by the operating system
    - Processes have more memory available than the machine has physically.
    - Physical memory divided into 4KB-sized pages and divided between processes.
    - The Virtual Memory Manager translates virtual addresses into physical addresses.
    - Pages not accessed by a process are written to disk and re-read if necessary: *swapping*
  - Virtual machines managed by a hypervisor
    - Several operating systems co-exist and seem to manage each their own computer, but physically there is only one single computer.
    - Physical devices (CPU, memory, harddisk and network adapter) shared between virtual machines.
    - A virtual machine contains a virtual CPU, virtual memory, a virtual disk and a virtual network adapter.
    - The hypervisor translates accesses by the virtual machine to these peripherals into accesses to the physical peripherals.

Virtualization

Without virtualization: One machine, one operating system

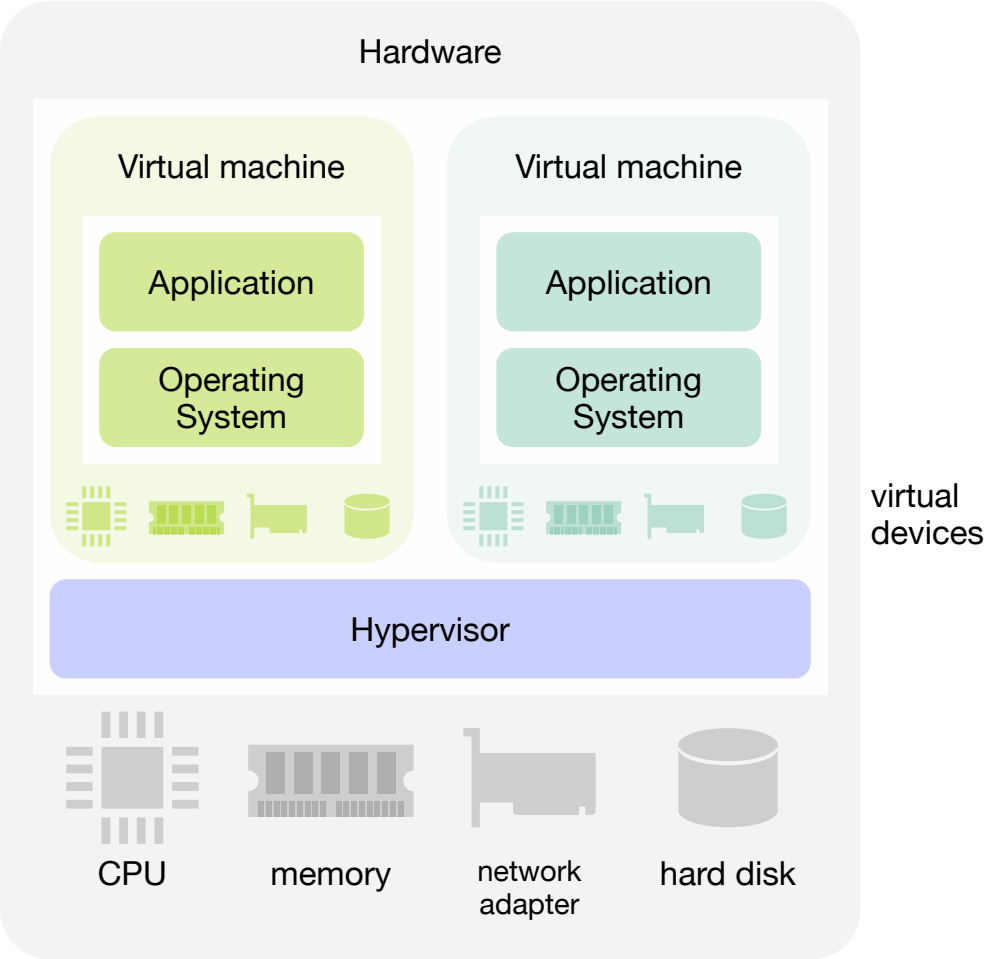
- The operating system manages the physical components of a computer.
  - CPU
  - Memory
  - Harddisk
  - Network adapter
  - ...



Virtualization

Subdivision of a physical machine into several virtual machines

- Virtualization of machines: creation of several virtual machines on a single physical machine.
- Physical devices are shared between virtual machines.
- Hypervisor translates accesses to virtual devices into accesses to physical devices.



Virtualization — Virtual machines

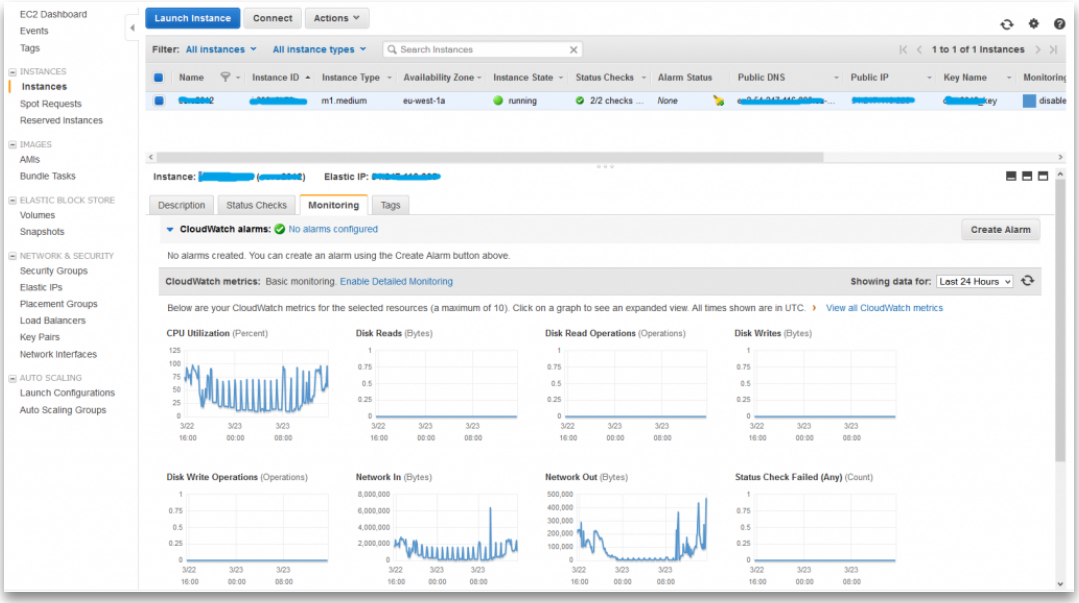
Properties

- **Binary compatibility**: Each machine thinks it owns the available hardware
  - No modification of the operating system or the applications
- **Interposition**:
  - All actions of a virtual machine have to go through the hypervisor.
- **Isolation**:
  - A program running in a virtual machine cannot access the data of another virtual machine.
    - **Software isolation**
    - **Fault isolation**
  - A virtual machine with high processing load cannot affect the performance of another virtual machine.
    - **Performance isolation**
- **Encapsulation**:
  - The complete state of a virtual machine can be captured in a file: *system image*
  - The file can be manipulated like any other file: transferred, duplicated, removed, ...

Amazon Web Services

Elastic Compute Cloud (EC2)

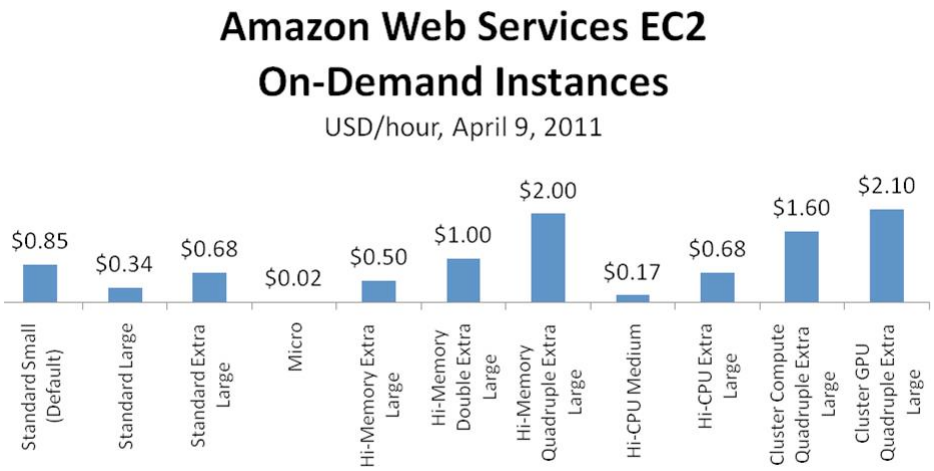
- Amazon EC2 presents a true virtual computing environment, allowing you to:
  - Use a web interface to launch instances with a variety of operating systems
    - which are bundled into *Amazon Machine Images* (AMI)
  - Load your instances with your custom application environment
  - Manage your network's access permissions
- Amazon EC2 reduces the time required to obtain and boot new server instances to minutes
  - This allows you to quickly scale capacity as your computing requirements change

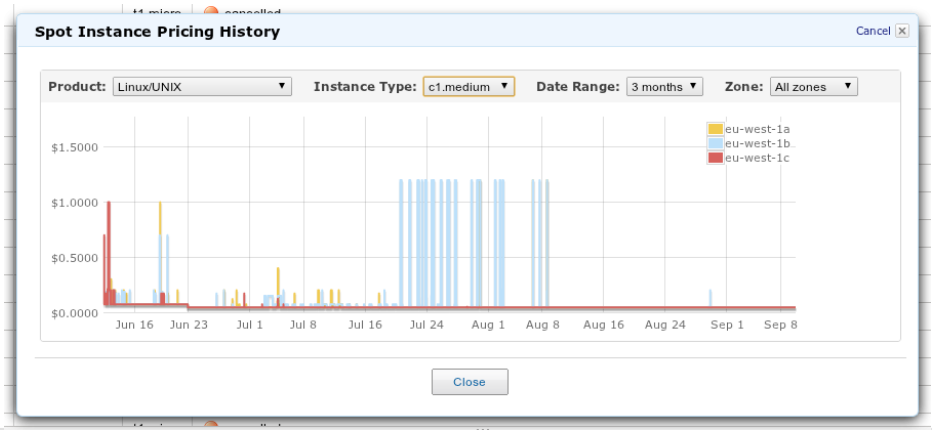


Amazon Web Services

EC2 instance models

- Amazon offers three instance models that differ in their availability and their pricing structure
  - On-Demand Instances
    - Pay-by-the hour
    - Start and stop as you wish
  - Reserved Instances
    - Pay a yearly upfront fee and receive a discount on the hourly charge
    - Start and stop as you wish
  - Spot Instances
    - Bid for unused EC2 capacity
    - Mention your Spot Price and if the market rate is less than your Bid, you get your instance
    - Instance automatically terminates if your Spot Price becomes less than the current market rate

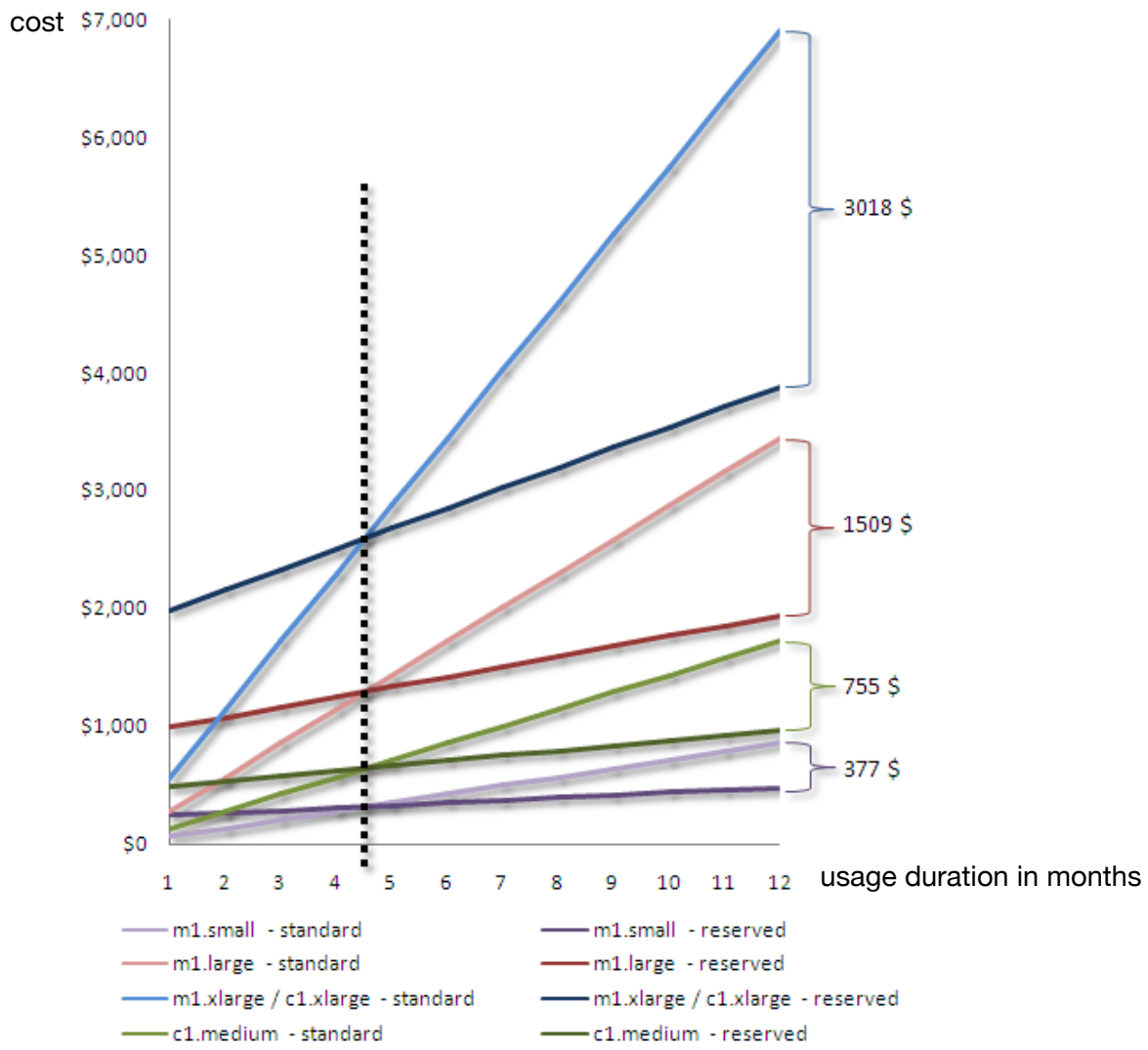




Amazon Web Services

EC2 on-demand vs. reserved instances — When are reserved instances cheaper?

- Example based on 2008 prices



Source: http://blog.guyegozy.com

Amazon Web Services

EC2 instance parameters

- CPU Power
  - Measured in *Elastic Compute Unit* (ECU) – Defined by Amazon as the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron/Zeon processor
- Memory
  - Measured in GB
- I/O performance
  - Three tiers: Low/Moderate/High
  - High-end instances have 10 Gigabit Ethernet

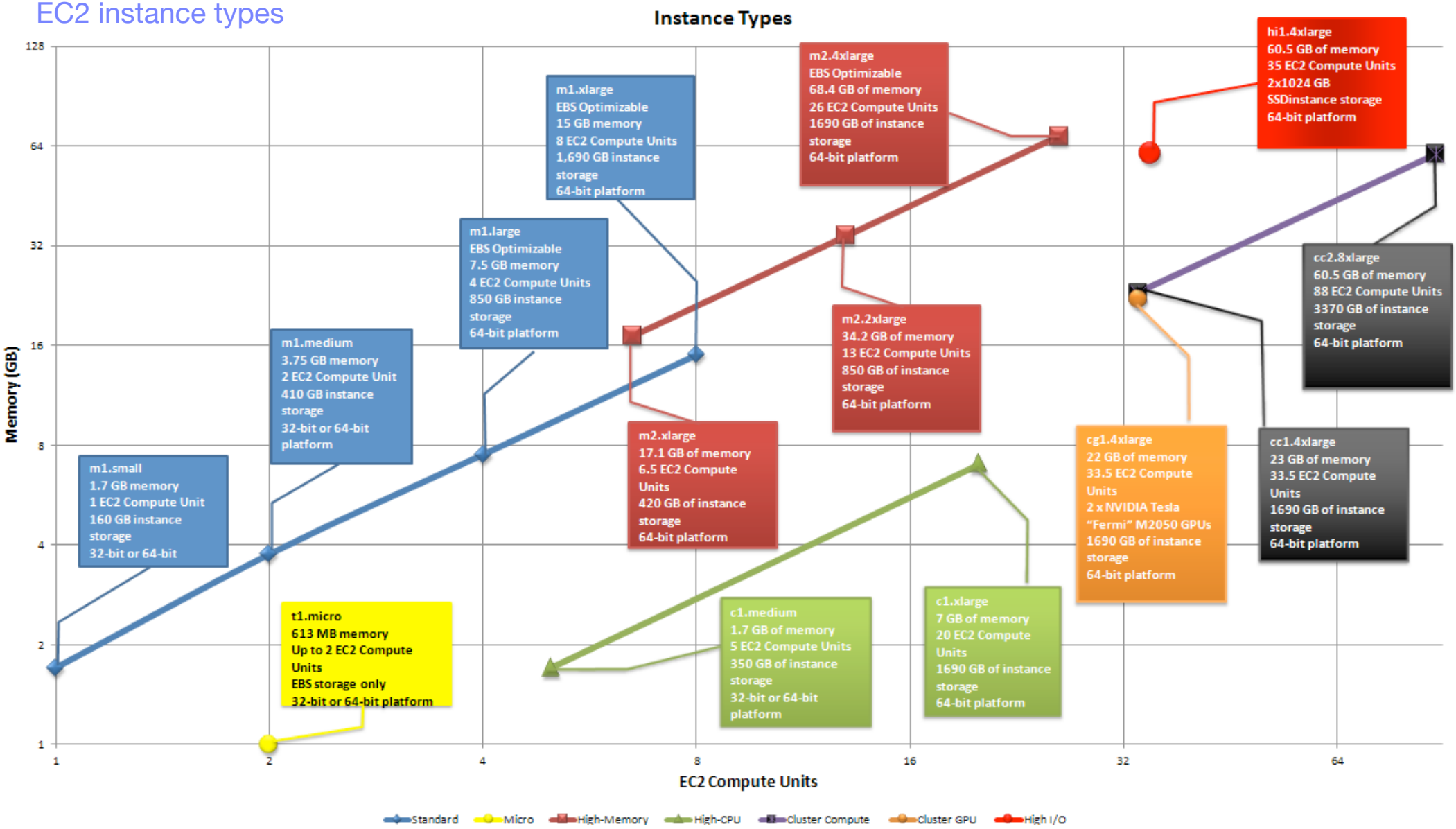
Amazon Web Services

EC2 instance types

Instance Name	Mem (GB)	CPU Capacity	Disk (GB)	Platform	On-Demand Pricing / hour (Linux)
Micro	0.59	Upto 2 ECUs	--	32/64	$0.02
Small	1.7	1 core - 1 ECU	160	32	$0.085
Large	7.5	2 cores, 2 ECUs each	850	64	$0.34
Extra Large	15	4 cores, 2 ECUs each	1690	64	$0.68
High-Mem Extra Large	17.1	2 cores, 3.25 ECUs each	420	64	$0.50
High-Mem Double Extra Large	34.2	4 cores, 3.25 ECUs each	850	64	$1.00
High-Mem Quad Extra Large	68.4	8 cores, 3.25 ECUs each	1690	64	$2.00
High CPU Medium	1.7	2 cores, 2.5 ECUs each	350	32	$0.17
High CPU Extra Large	7	8 cores, 2.5 ECUs each	1690	64	$0.68
Cluster Compute Quad XL	23	33.5 ECUs	1690	64	$1.30
Cluster GPU Quad XL	22	33.5 ECUS + 2x GPUs	1690	64	$2.10

Amazon Web Services

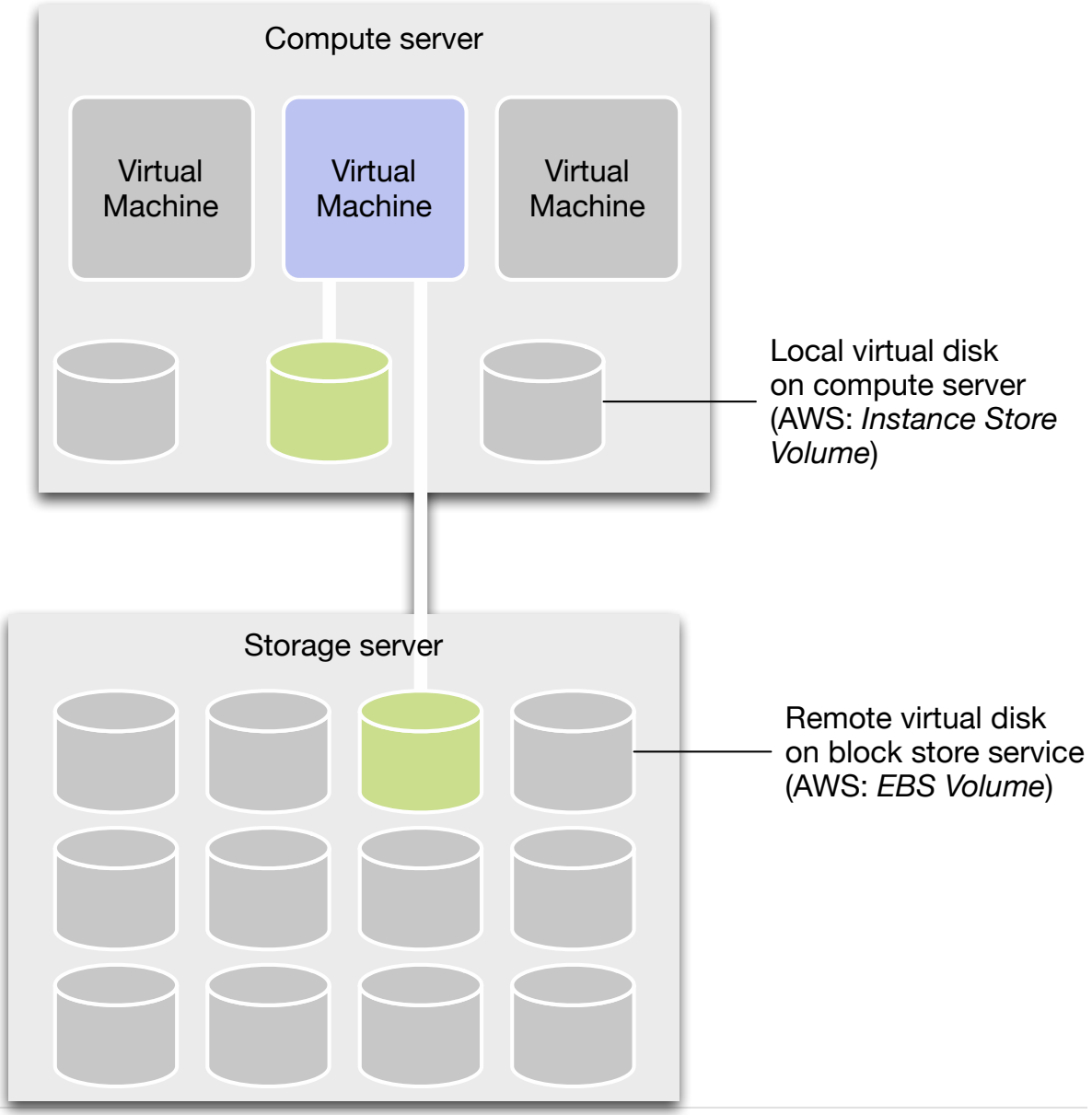
EC2 instance types



Amazon Web Services

Virtual machines and virtual disks

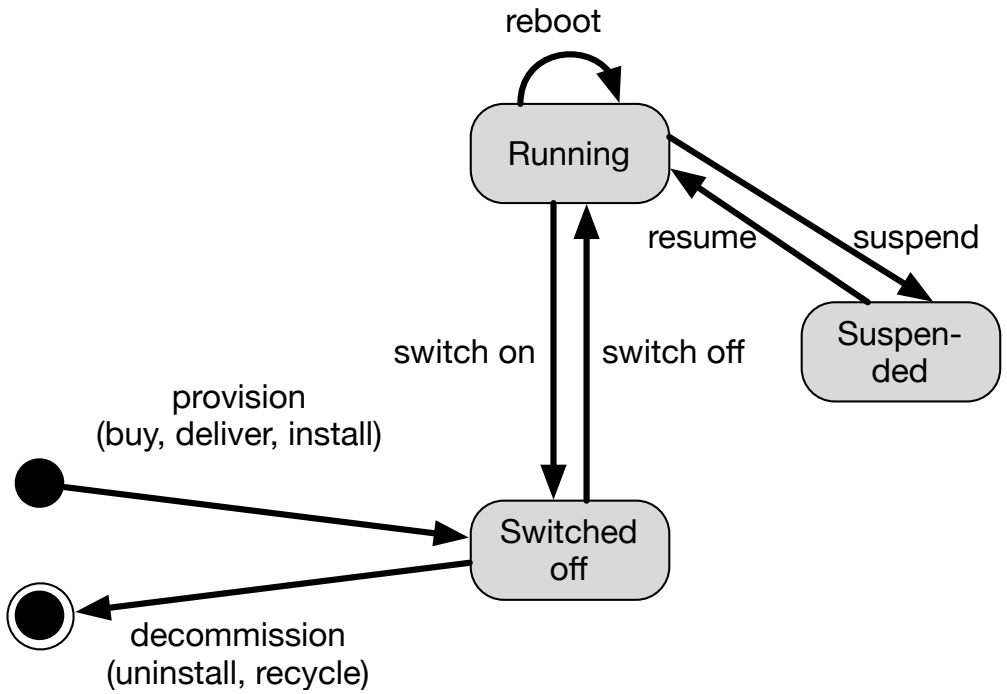
- A virtual machine can have one or more virtual disks attached to it
- There are two types of virtual disks:
  - Virtual disk managed by the storage service and allocated on a storage server: ***EBS Volume***
    - Lifecycle independent of virtual machine
  - (Rarely used nowadays:) Virtual disk allocated on the same server that hosts the virtual machine: ***Instance Store Volume***
    - The disk is deallocated when the virtual machine is deallocated



Amazon Web Services

Life cycle of a physical server

- A physical server can be switched on and off, suspended and resumed.

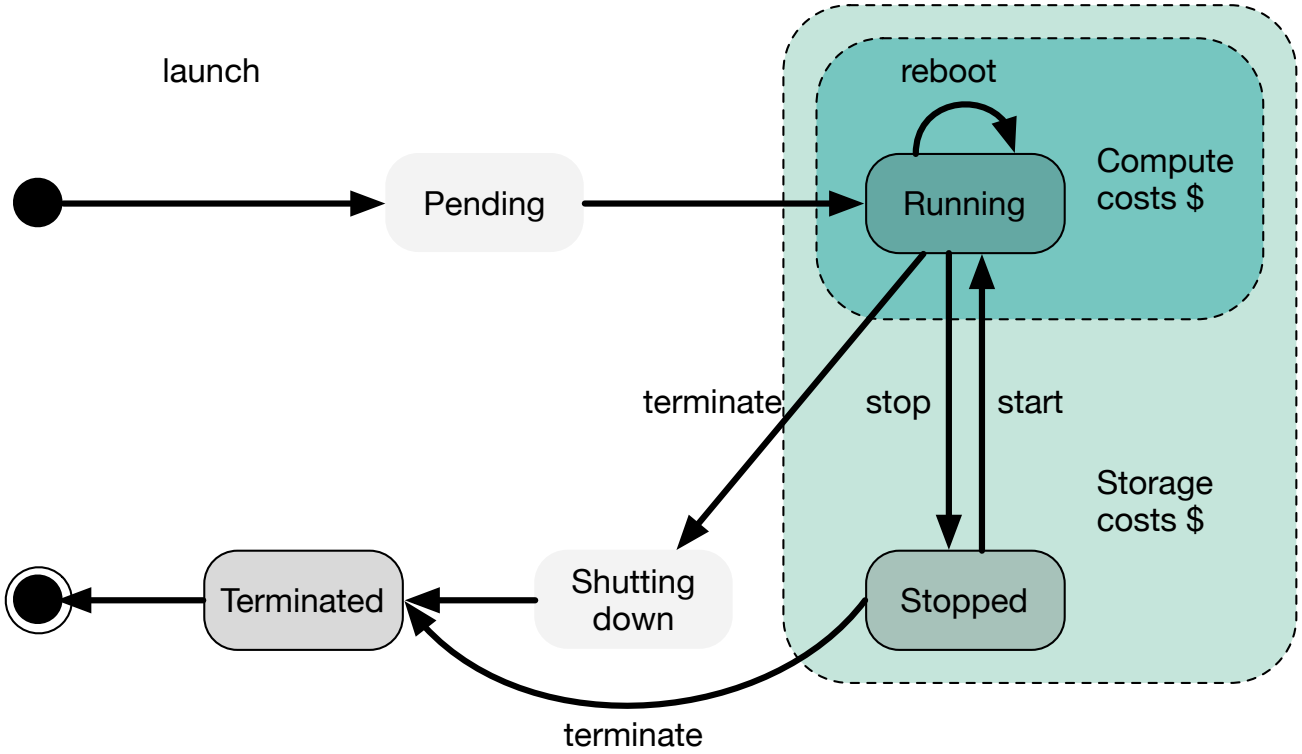


- An EC2 instance cannot be suspended, only "switched off” (stopped).

Amazon Web Services

Life cycle of an EC2 Instance

- EC2 Instance with an EBS Volume with “Delete on Termination” enabled (this is the default case)

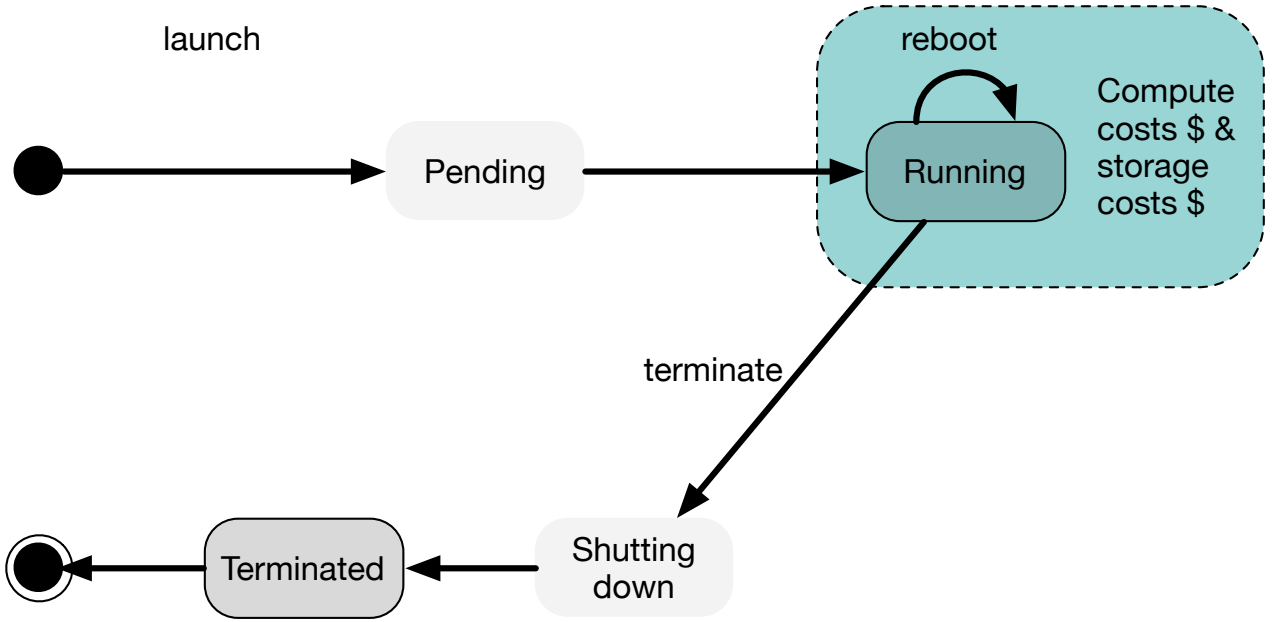


- When the instance is terminated, the disk is deallocated as well and the data is lost.
- To keep the data, “Delete on Termination” has to be disabled for the disk ($).

Amazon Web Services

Life cycle of an EC2 Instance

- EC2 Instance with an instance store volume

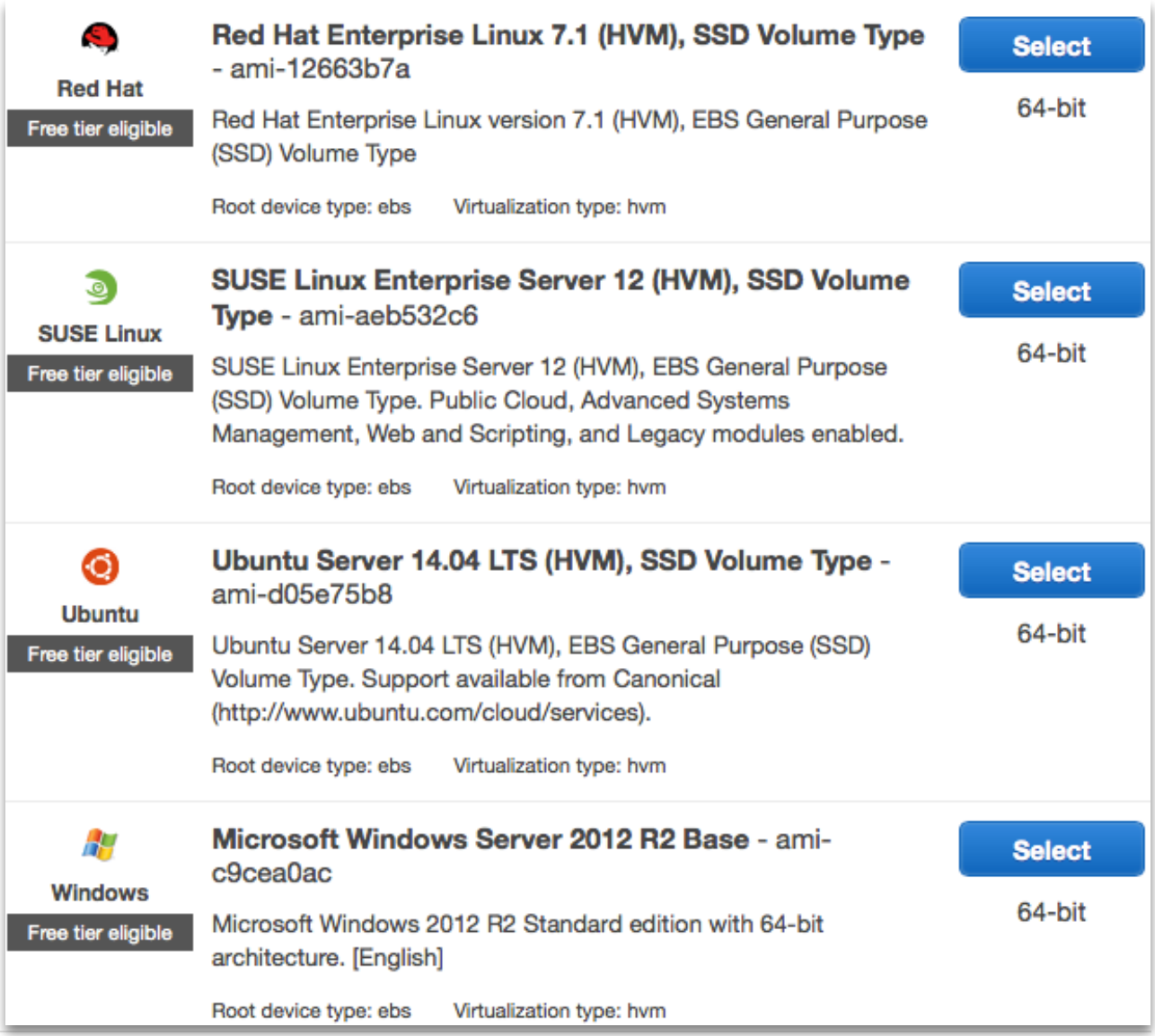


- When the instance is terminated, the disk is deallocated as well and the data is lost.
- In this case no way to keep the data beyond the lifetime of the instance.

Virtual machine Images

Amazon Machine Images (AMI)

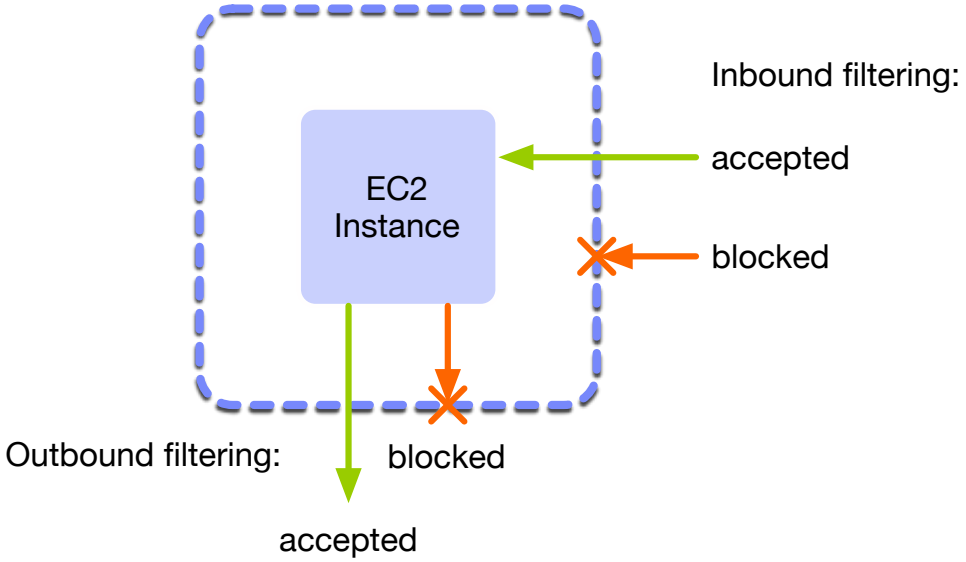
- When creating a virtual machine, the operating system is already installed.
  - AWS copies a virtual machine image *(Amazon Machine Image, AMI)*
- Thousands of AMIs available, some provided by Amazon, some by the community.
  - All major operating systems available.
- An AMI may also contain a complete software stack with middleware and applications.
- User is able to create his own AMIs.



Amazon Web Services

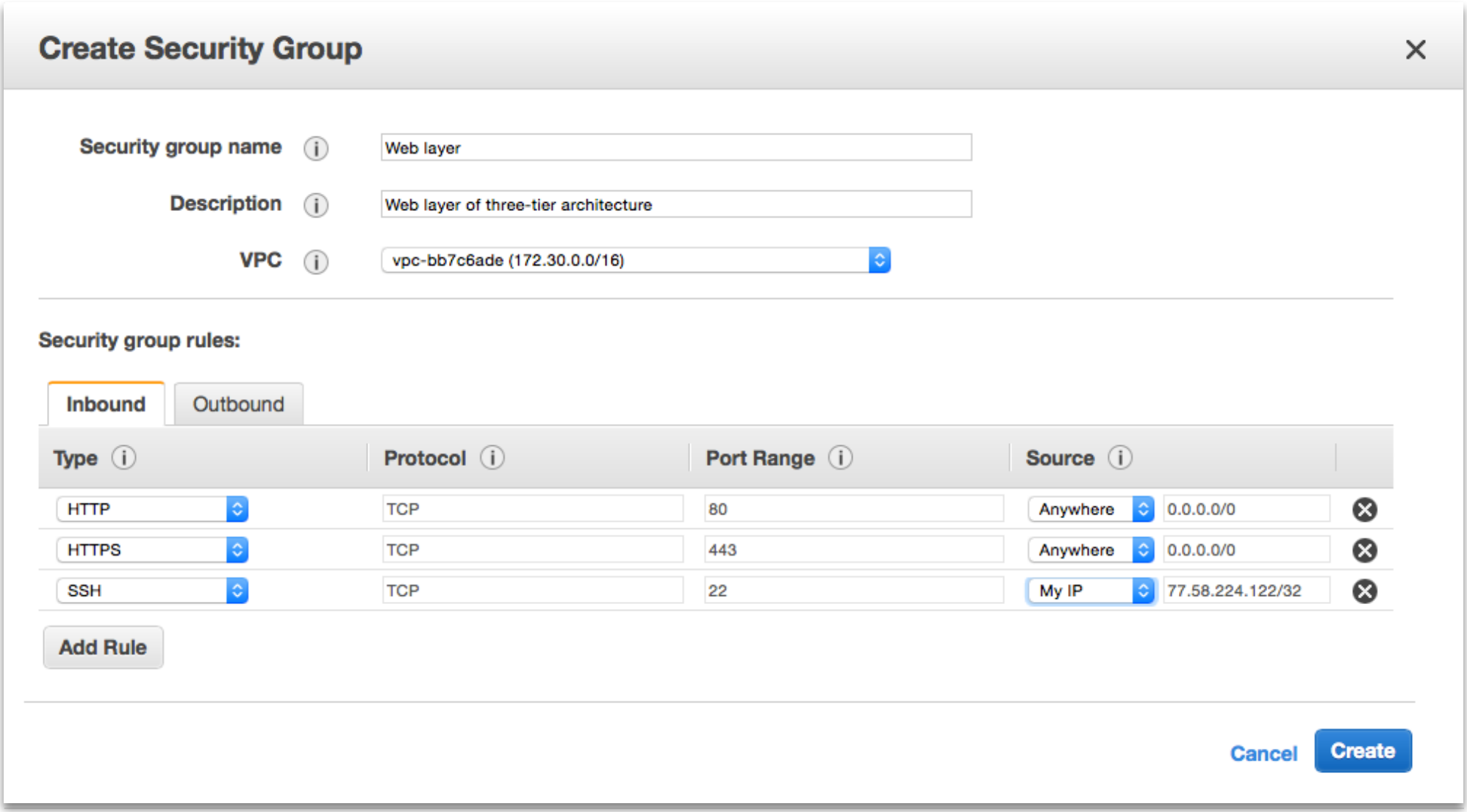
Firewalls / Security Groups

- An EC2 instance has a public IP address and is therefore accessible from the Internet.
  - This is a security risk.
- Each virtual machine comes with a mandatory virtual firewall.
- The configuration of a firewall is called *Security Group.*
  - A configuration can be reused for several firewalls.
- Firewall performs inbound and outbound filtering.
  - Based on protocols / port numbers
  - Based on IP addresses
- To be able to log into a Linux instance, port 22 (SSH) has to be open.



Amazon Web Services

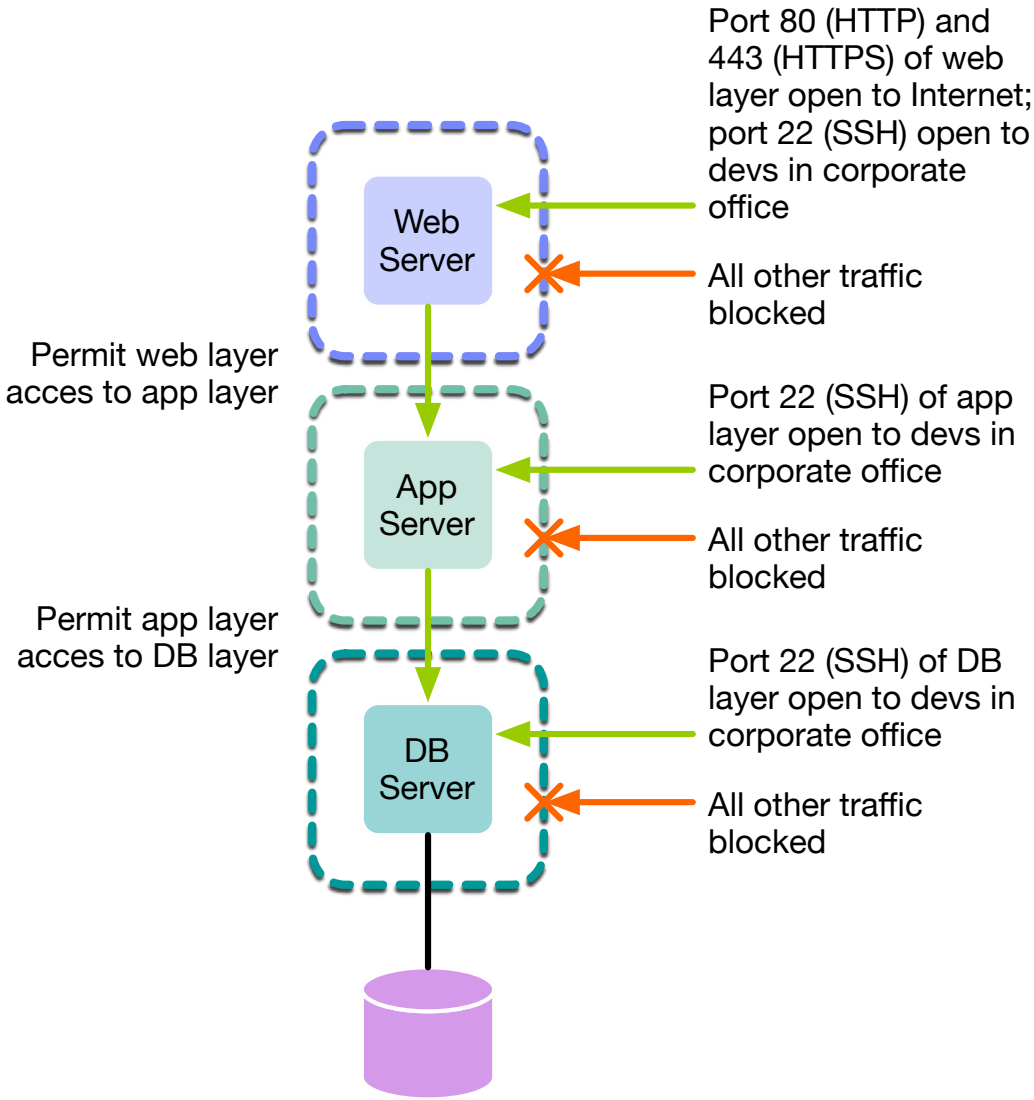
Creation of a Security Group



Amazon Web Services

EC2 Security Groups — Example configuration

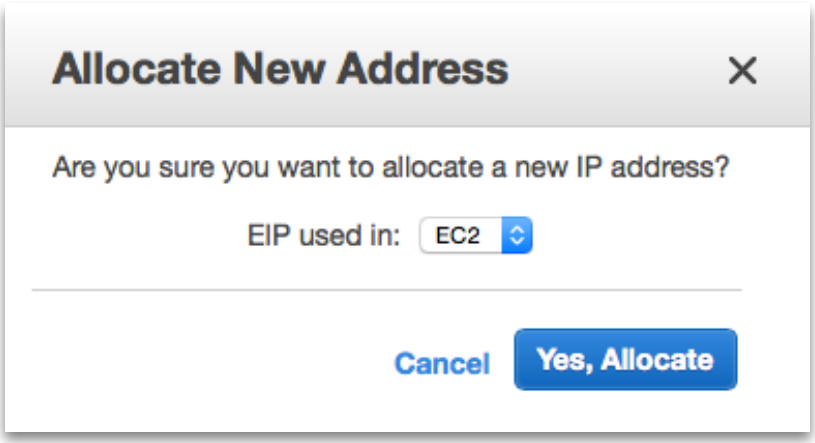
- Three-tier web application
- Restrictive configuration
- Distinguish access from
  - open Internet: HTTP and HTTPS open
  - corporate office (developers): SSH open
  - another layer of the web application: specific ports open

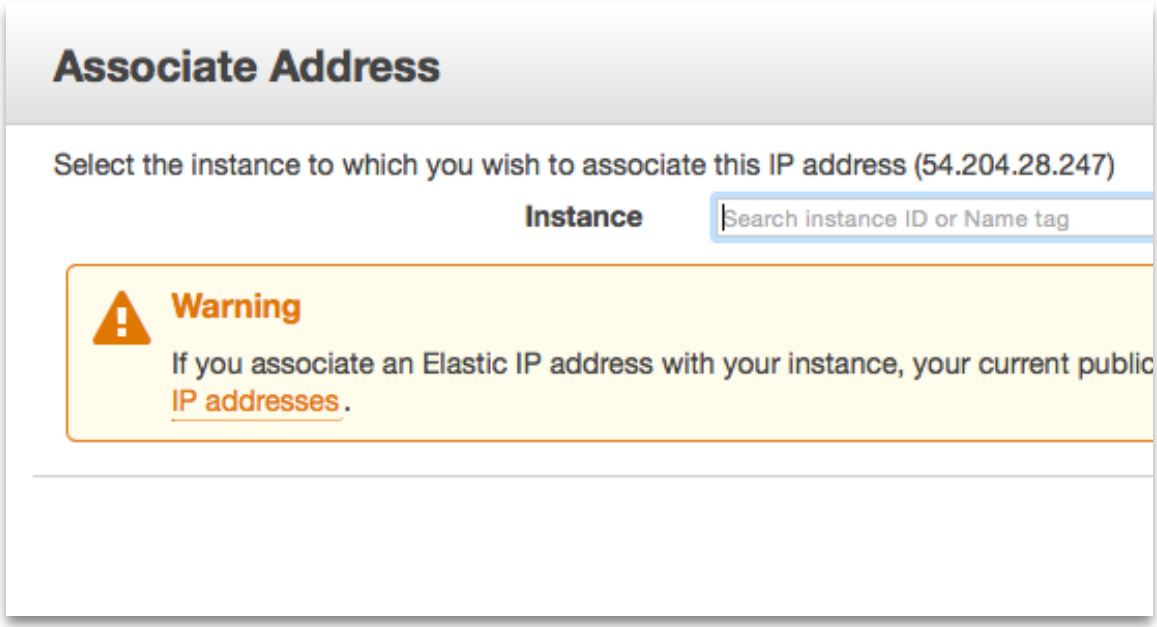


Amazon Web Services

EC2 Elastic IP Address

- An EC2 Elastic IP Address is a fixed IP address.
- It can be assigned to any instance in a region.
- Allows to keep the address when changing the instance.

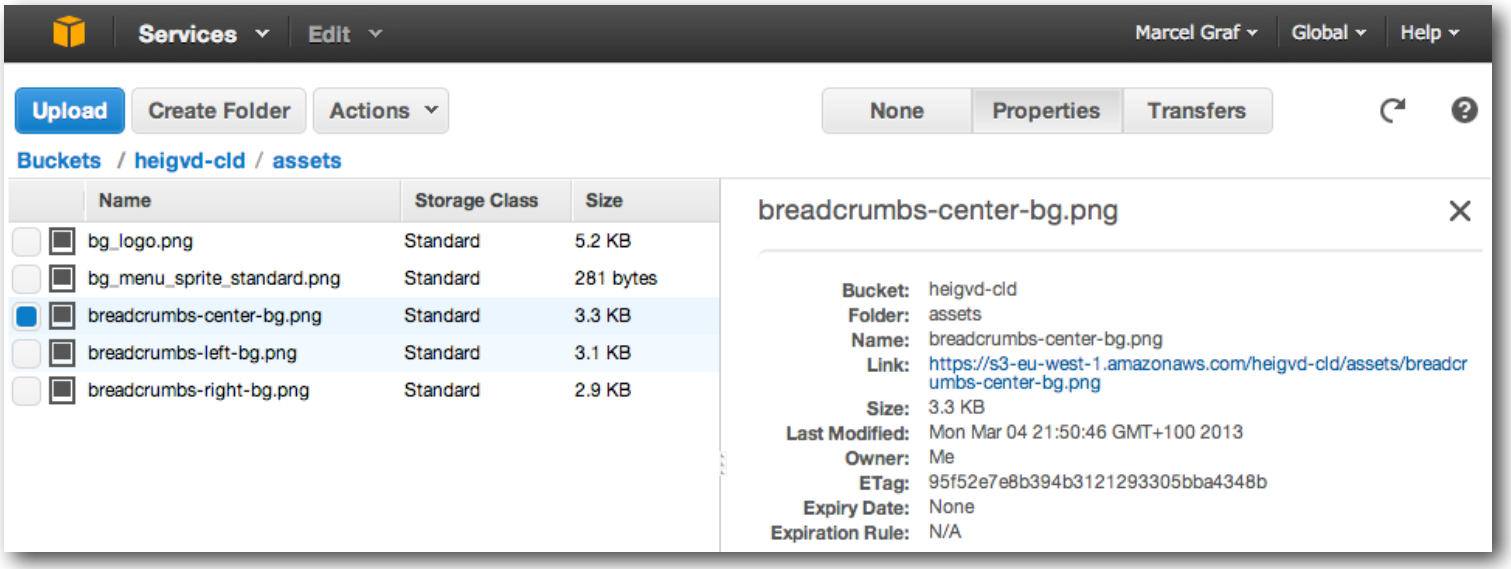




Amazon Web Services

Simple Storage Service (S3)

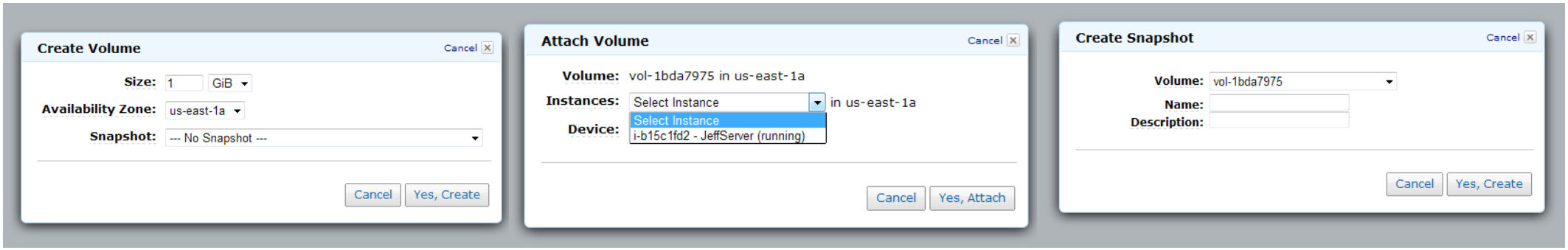
- Amazon Simple Storage Service (S3) is a high-performance, highly-available web-oriented storage service that supports very large files
- You can write, read and delete *objects* (files) into S3 containing from 1 byte to 5 TB of data. The number of objects that can be stored is unlimited.
- Each object is stored in a *bucket* and a bucket can be stored in one of several regions.
- Objects stored in a region never leave the region unless you transfer them out.



Amazon Web Services

Amazon Elastic Block Store (EBS)

- Amazon Elastic Block Store (EBS) offers persistent storage for EC2 instances
  - EBS volumes provide off-instance storage that persists independently from the life of an instance
- EBS provides the ability to create point-in-time consistent snapshots of your volumes that are then stored in S3, and automatically replicated across multiple available zones
- These snapshots:
  - Can be used as the starting point for new EBS volumes
  - Can protect your data for long term durability
  - Can be easily shared with co-workers and other AWS developers



Amazon Web Services
Other products / services / features

Amazon Web Services

Compute

EC2
Virtual Servers in the Cloud

EC2 Container Service
Run and Manage Docker Containers

Elastic Beanstalk
Run and Manage Web Apps

Lambda
Run Code in Response to Events

Storage & Content Delivery

S3
Scalable Storage in the Cloud

CloudFront
Global Content Delivery Network

Elastic File System PREVIEW
Fully Managed File System for EC2

Glacier
Archive Storage in the Cloud

Storage Gateway
Integrates On-Premises IT Environments with Cloud Storage

Database

RDS
MySQL, Postgres, Oracle, SQL Server, and Amazon Aurora

DynamoDB
Predictable and Scalable NoSQL Data Store

ElastiCache
In-Memory Cache

Redshift
Managed Petabyte-Scale Data Warehouse Service

Networking

VPC
Isolated Cloud Resources

Direct Connect
Dedicated Network Connection to AWS

Route 53
Scalable DNS and Domain Name Registration

Developer Tools

CodeCommit
Store Code in Private Git Repositories

CodeDeploy
Automate Code Deployments

CodePipeline
Release Software using Continuous Delivery

Management Tools

CloudWatch
Monitor Resources and Applications

CloudFormation
Create and Manage Resources with Templates

CloudTrail
Track User Activity and API Usage

Config
Track Resource Inventory and Changes

OpsWorks
Automate Operations with Chef

Service Catalog
Create and Use Standardized Products

Security & Identity

Identity & Access Management
Manage User Access and Encryption Keys

Directory Service
Host and Manage Active Directory

Trusted Advisor
Optimize Performance and Security

Analytics

EMR
Managed Hadoop Framework

Data Pipeline
Orchestration for Data-Driven Workflows

Kinesis
Real-time Processing of Streaming Big Data

Machine Learning
Build Smart Applications Quickly and Easily

Mobile Services

Cognito
User Identity and App Data Synchronization

Device Farm
Test Android, Fire OS, and iOS apps on real devices in the Cloud

Mobile Analytics
Collect, View and Export App Analytics

SNS
Push Notification Service

Application Services

API Gateway
Build, Deploy and Manage APIs

AppStream
Low Latency Application Streaming

CloudSearch
Managed Search Service

Elastic Transcoder
Easy-to-use Scalable Media Transcoding

SES
Email Sending Service

SQS
Message Queue Service

SWF
Workflow Service for Coordinating Application Components

Enterprise Applications

WorkSpaces
Desktops in the Cloud

WorkDocs
Secure Enterprise Storage and Sharing Service

WorkMail PREVIEW
Secure Email and Calendaring Service

Amazon Web Services

Cost comparison: do-it-yourself vs. AWS

Annual Cost Comparison (100% utilization)				
	Do-It-Yourself	**EC2 On-Demand**	**EC2 Reserved (1 Year Term)**	**EC2 Reserved (3 Year Term)**
Usage Costs	-	$ 157,680	$ 75,411	$ 48,123
Server Hardware	$ 20,129	-	-	-
Network Hardware	$ 4,026	-	-	-
Hardware Maintenance	$ 28,986	-	-	-
Operating System	$ –	-	-	-
Facility Expense	$ 131,382	-	-	-
Remote Hands Support	$ 1,014	-	-	-
Data Transfer Costs	$ 10,071	$ 6,138	$ 6,138	$ 6,138
TOTAL COST	**$ 195,608.00**	**$ 163,818.00**	**$ 81,550.00**	**$ 54,263.00**

AWS Terminology

AWS term	Generic term	AWS term	Generic term
EC2	IaaS offering	Elastic Block Store (EBS) volume	Virtual disk on a SAN
EC2 Instance	Virtual machine	Instance store volume	Virtual disk co-located with virtual machine
Amazon Machine Image (AMI)	Virtual machine image	S3	Object storage service
Security Group	Firewall configuration	Cloud Watch	Monitoring service
Elastic IP address	Static external IP address		
Volume	Virtual disk		

See also: Amazon Web Services in Plain English, https://www.expeditedssl.com/aws-in-plain-english