

“Alexandru Ioan Cuza” University of Iasi
Faculty of Computer Science

Master
Course
I+II

Conf. Dr. Lenuța Alboai
adria@info.uaic.ro



Cloud Computing - Properties and characteristics

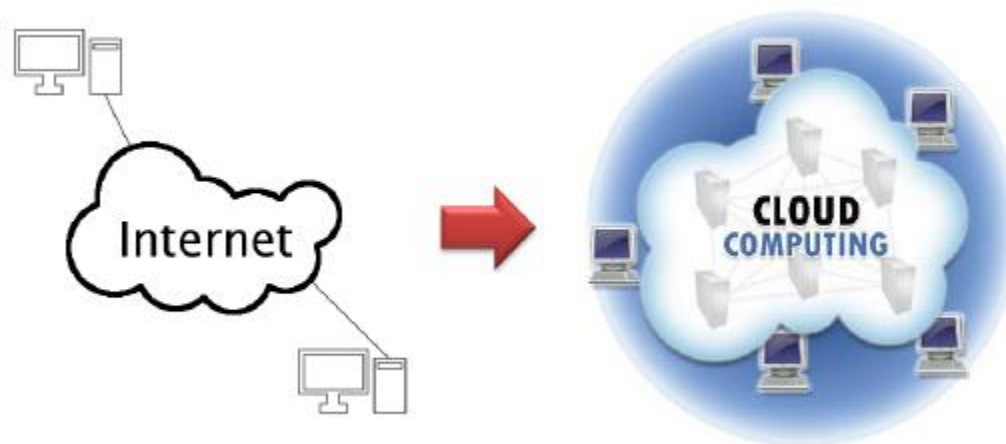
Content

- Cloud Computing?
 - Definitions
 - Characteristics
- Services in cloud
- Deployment models

Cloud Computing | Definitions

Whatis.com

- *“The name cloud computing was inspired by the cloud symbol that's often used to represent the Internet in flowcharts and diagrams. Cloud computing is a general term for anything that involves delivering hosted services over the Internet.”*



Cloud Computing | Definitions

Wikipedia

- *“Cloud computing is Internet-based computing, whereby shared resources, software, and information are provided to computers and other devices on demand, like the electricity grid.”*



- *“Cloud computing is a style of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet.”*

Cloud Computing | Definitions

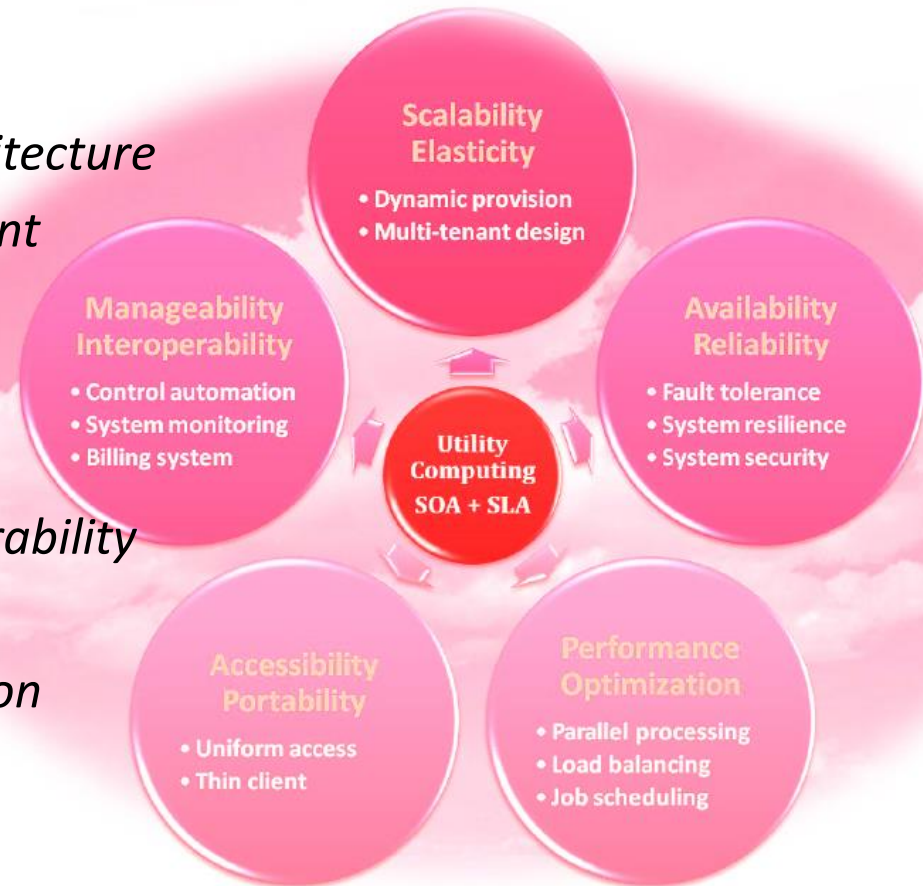
Buyya



- *“A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers.”*₅

Cloud Computing

- Key aspects
 - *Utility Computing*
 - *SOA – Service Oriented Architecture*
 - *SLA – Service Level Agreement*
- Proprieties and characteristics:
 - *scalability and elasticity*
 - *availability and reliability*
 - *manageability and interoperability*
 - *accessibility and portability*
 - *performance and optimization*
- Used technics:
 - Virtualization
 - Parallel and distributed computing
 - Web Services




Cloud Computing

Utility
Computing
SOA + SLA

- **Utility Computing** (Course 2)
 -
 - Example:
 - 2006, Sun Grid Compute Utility
 - Offer computing power in pay-per-use manner (1\$/CPU/hour)
 - 2009 , Sun: Open Cloud Platform
 - Sun Cloud Storage Service and Sun Cloud Compute Service
- *utility computing* can be successful if ... => there is an accessible interface that can be easily understood and explored by developers

Cloud Computing



Utility
Computing
SOA + SLA

- What is a Web Service?

Definitions

- “Web service is self-describing and stateless modules that perform discrete units of work and are available over the network”
- “Web service providers offer APIs that enable developers to exploit functionality over the Internet, rather than delivering full-blown applications”
- **SOA - Service Oriented Architecture**
 - A service collection that communicates
 - Contains design principles that are used in development and integration

Cloud Computing

Utility
Computing
SOA + SLA

“A cloud needs an access API”

- An API enables the use of services
- Example:
 - Amazon’s EC2 API is an *SOAP API- and HTTP Query-based* , with operations such us: creation, storage, management for AMI (Amazon Machine Images)
 - Kenai Cloud API (Sun) is an REST API used to create and manage various resources in cloud (computing, storage, network components)
- Security mechanism are used for authorized access (e.g. Amazon – X.509)
- Obs. There are no standards for cloud APIs

Cloud Computing

Utility
Computing
SOA + SLA

- **QoS (Quality of Service)** *“is a set of technologies for managing network traffic in a cost effective manner to enhance user experiences for home and enterprise environments”* (definitions for a computer network)
- QoS includes:
 - End customer evaluations
 - Evaluations related to technical aspects (e.g. error rates, bandwidth, transmission delay, availability)
- **SLA – “Service-level agreement** *is a contract between a network service provider and a customer that specifies, usually in measurable terms (QoS), what services the network service provider will furnish.”*
 - There are metrics that assure performance (upload/download time, RTT, transfer rate)
 - Management details
 - Penalties if there were problems
 - Security aspects

Cloud Computing

Scalability
Elasticity

- Dynamic provision
- Multi-tenant design

Proprieties:

- *Scalability*
 - Represents a characteristic of a system, network or process which indicates the ability to easily handle the increase of data management
- *Elasticity*
 - Represents an infrastructure ability to adapt to the requirements based on real-time analysis methodologies of the entire system
 - *Elasticity* is an ability to increase the system performance when the demands are high and to decrease the performances when demands are low
- *How to get it?*
 - ***Dynamic Provisioning***
 - **multi-tenant design**

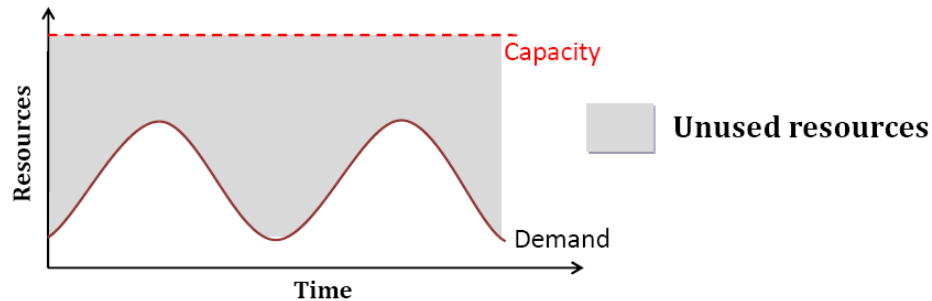
Cloud Computing

Scalability
Elasticity

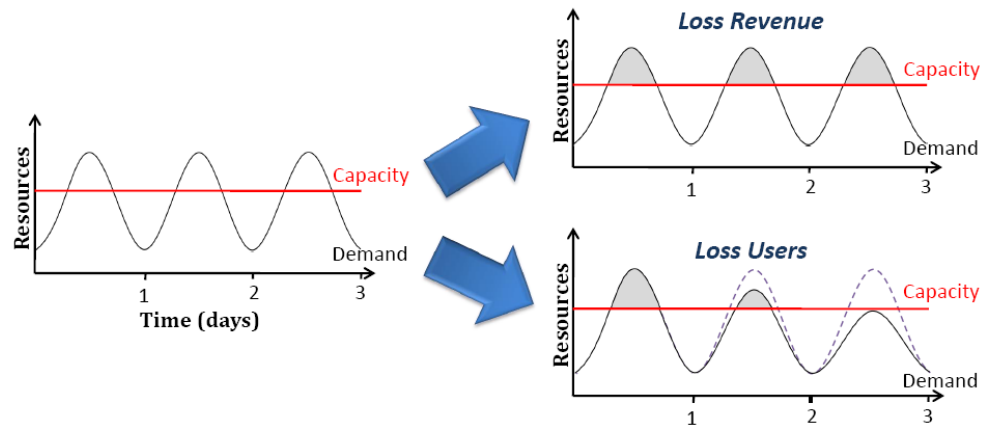
- Dynamic provision
- Multi-tenant design

Dynamic Provisioning:

- Traditional model
 - Problem 1: Overestimation of the used resource



- Problem 2: Underestimation of the used resources
Example: Twitter, LA Times etc...



Cloud Computing

Scalability
Elasticity

- Dynamic provision
- Multi-tenant design

Dynamic Provisioning

“A cloud needs elasticity: scaling your application as demand rises and falls.”

Examples from Amazon EC2: we want to assure scalability from 2 to 20 instances =>elasticity). If CPU usage exceeds 80 % from its capacity, a new instance is add; if it is lower than 40% for 10 minutes , an instance is removed

CreateLoadBalancer:

```
AvailabilityZones = us-east-1a
LoadBalancerName = MyLoadBalancer
Listeners = lb-port=80,instance-port=8080,protocol=HTTP
```

CreateLaunchConfiguration:

```
ImageId = myAMI
LaunchConfigurationName = MyLaunchConfiguration
InstanceType = m1.small
```

CreateAutoScalingGroup:

```
AutoScalingGroupName = MyAutoScalingGroup
AvailabilityZones = us-east-1a
LaunchConfigurationName = MyLaunchConfiguration
LoadBalancerNames = MyLoadBalancer
MaxSize = 20
MinSize = 2
```

CreateOrUpdateScalingTrigger:

```
AutoScalingGroupName = MyAutoScalingGroup
MeasureName = CPUUtilization
Statistic = Average
TriggerName = MyTrigger1a
Namespace = AWS/EC2
...
LowerThreshold = 40
LowerBreachScaleIncrement = -1
UpperThreshold = 80
UpperBreachScaleIncrement = 1
BreachDuration = 600
```

Cloud Computing

Scalability
Elasticity

- Dynamic provision
- Multi-tenant design

Multi-tenant design

- Principle in software architecture where a single software instance runs on a server and multiple clients are served (tenants)
 - In an multi-tenant architecture applications share in a virtual manner data and configurations => each client (e.g. company, end-user,...) use a personalized instance of a virtual application
 - Multi-tenant applications
 - are provided with a certain level of customization that can be adapted to customer requirements
 - can provide a certain level of security and resistance-> It facilitates the *release management* process
- ... Docker...Virtualization (next course)

Cloud Computing

Availability
Reliability

- Fault tolerance
- System resilience
- System security

Proprieties:

- *Availability*
 - The degree to which a system / subsystem equipment is in a functional state and is ready to work in any time
 - Cloud systems require high availability (~ 99.999%)
 - If the supplier is subject to a DOS attack, will client's critical systems collapse?

Service and Outage	Duration	Date
S3 outage: authentication service overload leading to unavailability	2 hours	2/15/08
S3 outage: Single bit error leading to gossip protocol blowup.	6-8 hours	7/20/08
AppEngine partial outage: programming error	5 hours	6/17/08
Gmail: site unavailable due to outage in contacts system	1.5 hours	8/11/08

Examples: AWS, AppEngine and Gmail indisponibility

[<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.htm>]

15

Cloud Computing

Availability
Reliability

- Fault tolerance
- System resilience
- System security

Proprieties:

- *Reliability*
 - The ability of a system or component to perform the duties required under specified conditions for a specified period of time.
- *How to obtain Availability and Reliability?*
 - *Fault tolerance*
 - *Resilience*
 - *Security*

Cloud Computing

Availability
Reliability

- Fault tolerance
- System resilience
- System security

Fault tolerance

- Fault tolerance is the property that enables a system to continue operating properly in the event of failure of some of its components.
- Measure the functionality level: the decrease or increase in a proportional degree with the failure severity (in naïve systems design even a small failure can collapse the entire application)

- Characteristics:

- There is no SPOF



- Preventing: The system should function during the repairing processes
- Error detection and isolation of the failed component
- Fault isolation to prevent the spread(*fault containment*)
- The existence of check points allowing restoration

Cloud Computing

Availability
Reliability

- Fault tolerance
- System resilience
- System security

Resilience

- The ability to provide and maintain an acceptable level of service in case of error in order to offer usual services
- Indicates the system's ability to return to the original state as quickly as possible after errors have occurred
- Events
 - Loss of power supply
 - Corrupted Database
 - Natural Disasters
- Strategies
 - Backup
 - Additional sources of energy, Uninterruptible Power Supply (UPS)
 - ...



Cloud Computing



“A cloud needs servers on a network, and they need a home “ -> Data center

Aspects:

- Structure: Layout, maintenance, physical and logical security
- Scalability provided at optimal costs
- Why?
 - 1.2% of US electricity consumption
 - Acquisition of hardware in huge quantities
 - =>The possibility of negotiation
 - Ex. Amazon: 90 million on 50,000 servers from Rackable / SGI in 2008 versus \$ 215 million
- Research for reduction of consumption(e.g. Google - *dynamic voltage/frequency scaling*)

Cloud Computing

– “A cloud needs servers on a network, and they need a home”



Photograph of Google's top-secret Dalles, OR data center, built near the Dalles Dam for access to cheap power. Note the large cooling towers on the end of each football-sized building on the left. These towers cool through evaporation rather than using more power-hungry chillers. Source: Melanie Conner, *New York Times*.

[The Cloud at Your Service, Jothy Rosenberg, Arthur Mateos,]
20

Cloud Computing

Availability
Reliability

- Fault tolerance
- System resilience
- System security

Cloud Computing Security

- It is a subfield of computer security and computer networks
- Reflects a broad set of policies, technologies and controls developed to protect data, applications and cloud infrastructure
- Aspects
 - Privacy:
 - Does Sensitive data remain confidential? (Is the Cloud provider honest?)
 - Integrity
 - How do I know cloud provider performs the correct calculation? Is my data stored without being altered?
 - Privacy
 - Cloud stores data from multiple clients and "anyone" can run data mining algorithms

Cloud Computing

Availability
Reliability

- Fault tolerance
- System resilience
- System security

Cloud Computing security

Challenges

- Increasing the area vulnerable to attack
 - Data is stored and handled outside the organization
 - Attackers can target communication link between the client and cloud provider
 - Cloud provider employees may be subject to phishing attacks
- Auditing
 - Estimation of risk, prevention, detection, response to attacks - is doable because the data is outside the organization
- Legal Issues and reliable transfer
 - Who is responsible for regulatory compliance?
 - What happens in the case of cloud services subcontracting ...?

Cloud Computing

Availability
Reliability

- Fault tolerance
- System resilience
- System security

Cloud Computing Security

- Cloud Computing is a **security nightmare** and it can't be handled in traditional ways (Course 13)



John Chambers
CISCO CEO

Cloud Computing

Manageability
Interoperability

- Control automation
- System monitoring
- Billing system

Proprieties:

- *Manageability*
 - Issues closely related to enterprise management features tailored to cloud computing systems
- **Interoperability**
 - A system propriety to own interfaces that are fully understood and that enable present and future interaction with other systems, without access restrictions or implementation limits

How to get it?

- *Control automation*
- System monitoring

Cloud Computing

Manageability
Interoperability

- Control automation
- System monitoring
- Billing system

Control automation and Monitoring

- *Autonomic Computing* – development of systems capable of self-management, to overcome management's increasing complexity of computing systems, now and in the future
- *Autonomic Components* – can be seen as some control entities: sensor (self-monitoring), self-regulating mechanisms, with some knowledge and planning abilities based on rules/policies; the action is performed according to its own condition and environment
- What can be monitored:
 - Status hardware level (physical and virtualized)
 - Measuring performance parameters resources
 - Network access patterns
 - System Logs

Cloud Computing

Manageability
Interoperability

- Control automation
- System monitoring
- Billing system

Billing

- Users pay what they use
- Cloud Service Provider
 - Monitors the system (previous slide)
 - Automatically computes the costs for used services and forwards the payment request



Cloud Computing

Performance Optimization

- Parallel processing
- Load balancing
- Job scheduling

Proprieties: Performance and Optimization

- Parallel processing
 - Hardware approaches (*multi-core* systems, *vector processor* systems, distributed systems(e.g. Cluster Computing, Grid Computing))
 - Software approaches (programming languages for parallel computing, platforms that provide automatic parallelization)
- Load balancing http://en.wikipedia.org/wiki/Parallel_computing#Types_of_parallelism
 - The technique implies a uniform distribution load on two or more computers, network links, CPUs, hard drives or other resources, in order to best use resources to maximize traffic to decrease response time and to avoid overloading
- Tasks planning
 - A job scheduler is a software system responsible for executions in the background (*~batch processing*)
 - Intensive computing tasks, tasks used in complex processing dependent on each other
 - Approaches: pre-defined workflows, automatic configuration system

Cloud Computing

Accessibility
Portability

- Uniform access
- Thin client

Proprieties

- Accessibility
 - Describes the level to which a product, device, service or environment is accessible to more customers
- Portability
 - It is the ability to access the service using any device, anywhere, continuously and in an adaptive manner to the resource availability
- *How can you get it?*
 - Uniform access
 - *Thin client*

Cloud Computing

Accessibility
Portability

- Uniform access
- Thin client

Uniform access

- Using different OS, different platforms should provide access to cloud services



Thin client

- A device or program that depends on another computer system for performing various computational roles
- Versus *fat client*
- Features
 - Devices at affordable prices
 - A greater variety of end devices
 - Simplification at the client level

Cloud Computing | Services



Cloud Computing | Services

Question: Would you move to a town and look for solutions to live?

- Would you build a new house?



- Would you buy a cold dark shell house?



- Would you live in a hotel?



Cloud Computing | Servicii

Question: How to build an IT department?

– IaaS (Infrastructure as a Service)

- You rent a virtualized infrastructure and you can build an IT system that you can fully control



– PaaS (Platform as a Service)

- You develop an IT system on an existing cloud platform, without care of low level resources



– SaaS (Software as a Service)

- You use the existing IT systems offered by a cloud provider and you are not aware of technical aspects



Cloud Computing

General architecture

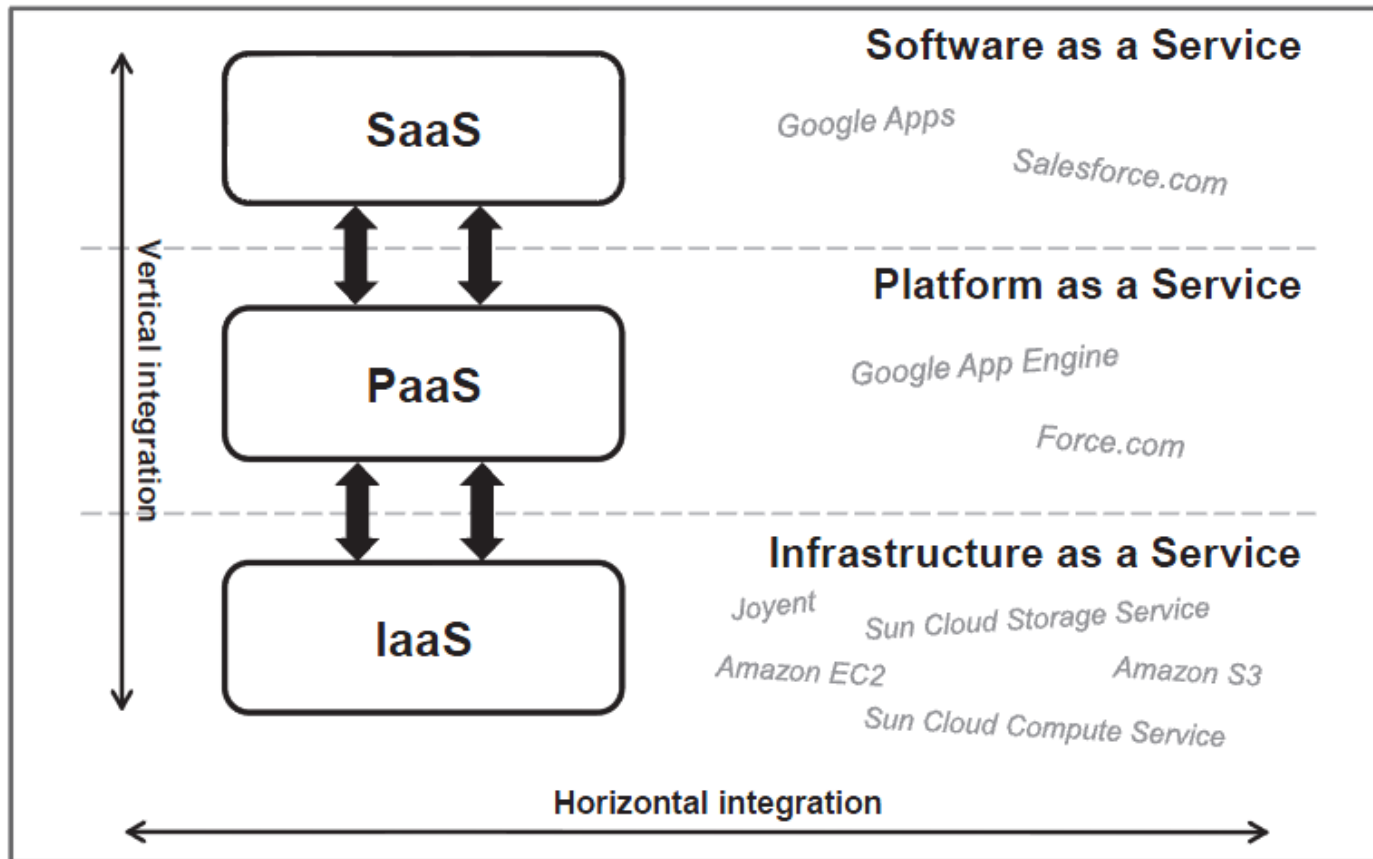
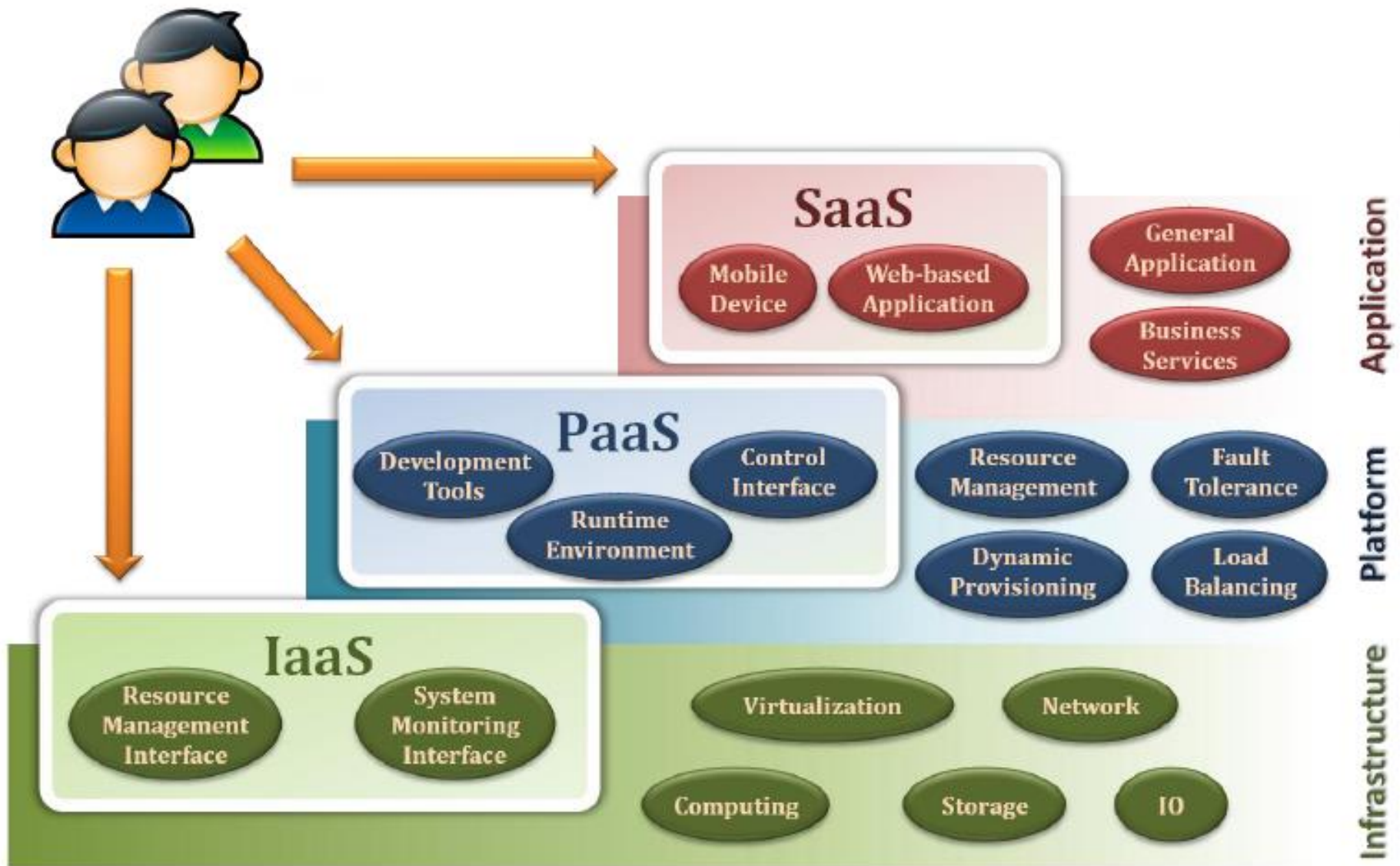


Figure: First IaaS, PaaS, SaaS providers

Cloud Computing | Services



Cloud Computing

Cloud Computing - Typical Architecture

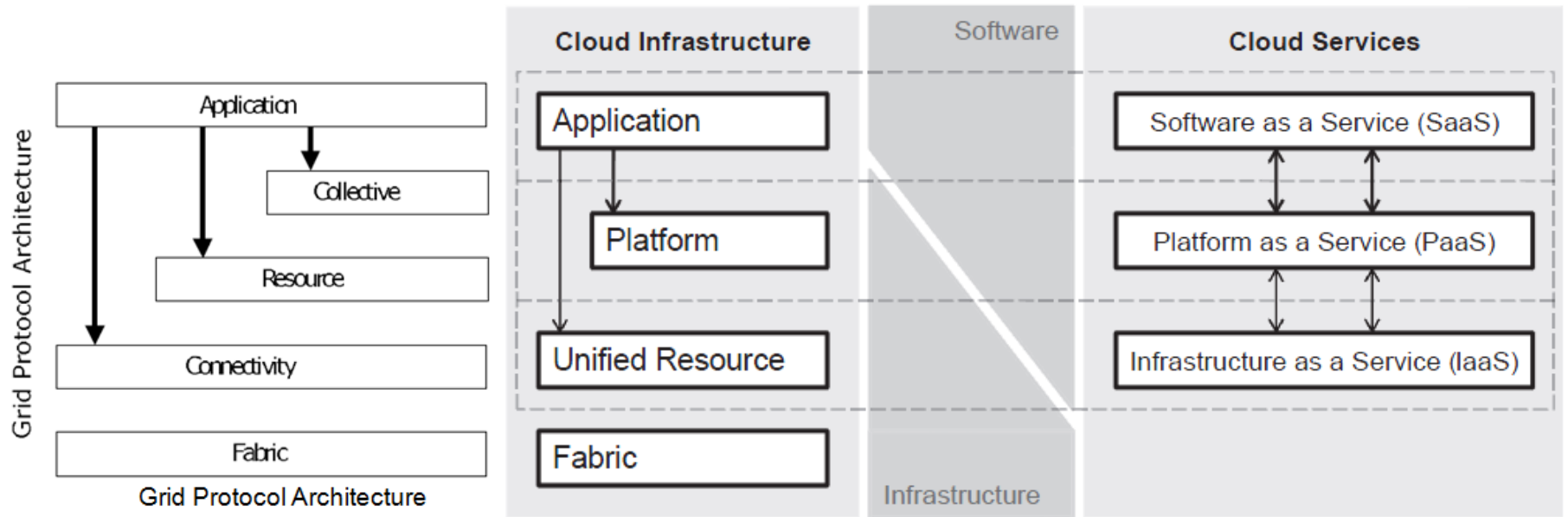


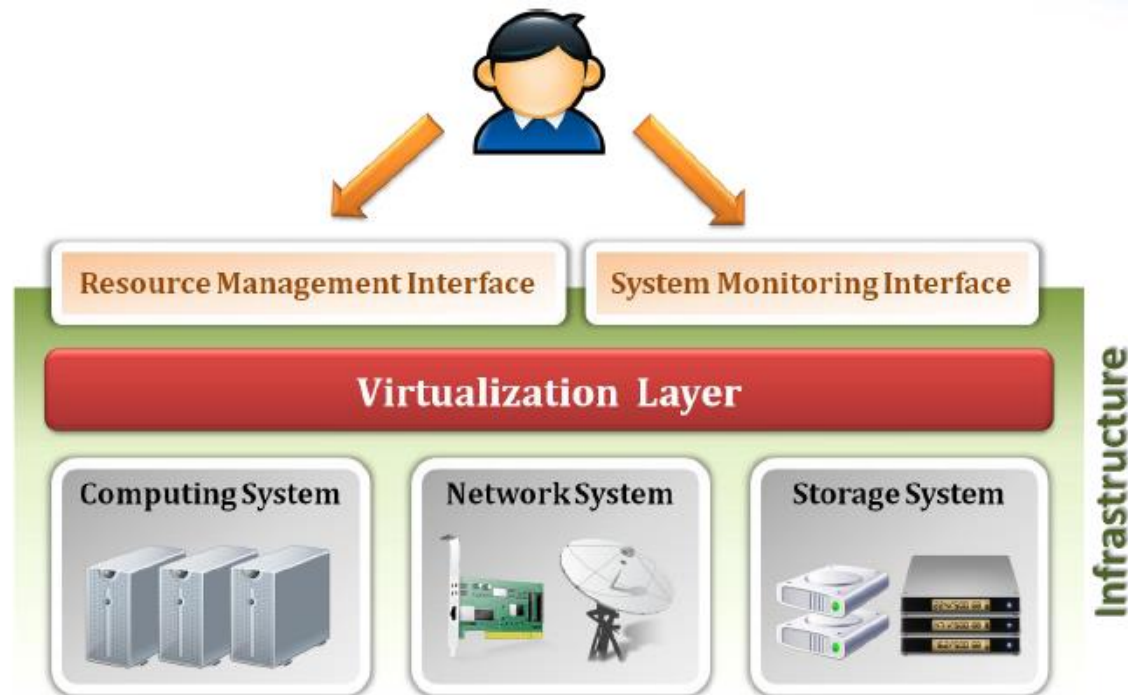
Figure: Cloud Architecture and Cloud Services (Foster, 2008)

Cloud Computing | Services

IaaS - Infrastructure as a Service

Note: *Fabric or Hosting Platform* - Provides the physical machine, operating system, the communication network, storage systems and virtualization software

- **IaaS** abstracts Fabric level and offer virtualized infrastructure (versus *raw hardware*) as a service , for processing, storage and communication



Cloud Computing | Services



IaaS - Infrastructure as a Service

- The consumer has no access to the fabric, but has control over operating systems, storage, application development and configurations related to network
- Example:
 - Amazon
 - Elastic Compute Cloud (EC2) for processing
 - Simple Storage Service (S3) for storage
 - Eucalyptus: Cloud open source implementation compatible with EC2 (used for on-premise (private) IaaS infrastructure)
 - OpenStack
 - ...

Cloud Computing | Services



IaaS - Infrastructure as a Service

- Key technology: virtualization recently -> Dockers
- Virtualization - is the emulation of one or more workstations within a single physical computer
- The hardware resources are transformed and virtualized (CPU, RAM, hard disk, network controller) => functional virtual machines that can run its own OSs

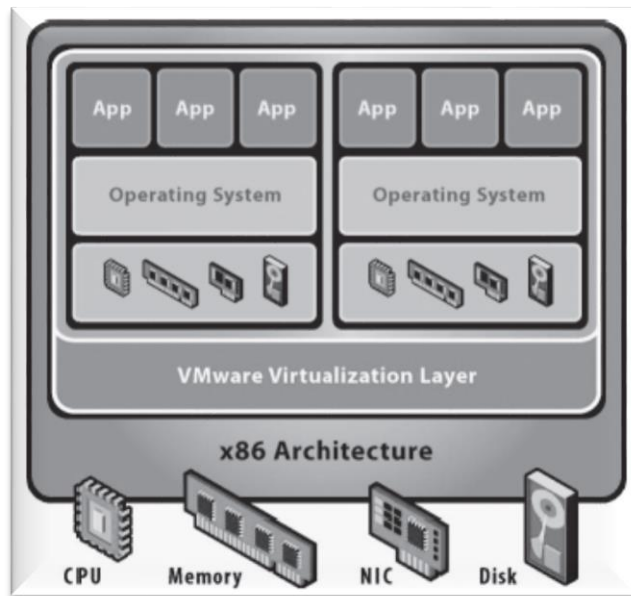


Figure. Architecture of virtual machines using VMware on x86

- 1999 VMware introduced the first virtualization application for x86 systems

[Cloud Computing
Virtualization
Specialist Complete]

Cloud Computing | Services

IaaS - Infrastructure as a Service

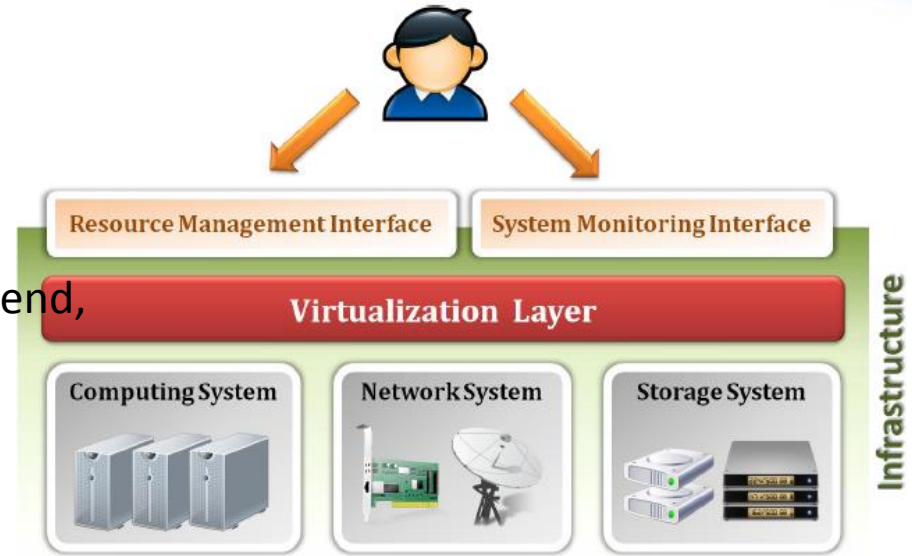
IaaS provider can offer:

RMI (Resource Management Interface)

- *Virtual Machine* – operations: create, suspend, reboot
- *Virtual Storage* – operations: allocation of space, free space, writing or reading
- *Virtual Network* – operations: allocation of IP addresses, domain registration, establish connections etc.

SMI (System Monitoring Interface) – examples of metrics for monitoring::

- *Virtual Machine*: CPU load, memory usage, IO load, internal network load.
- *Virtual Storage*: using virtual space, duplication, bandwidth to access the storage device
- *Virtual Network*: virtual network bandwidth, level of network load



Cloud Computing | Services

IaaS - Infrastructure as a Service

Virtualization supports properties as: *scalability and elasticity*

- *availability and reliability*
- *manageability and interoperability*



Cloud Computing | Services

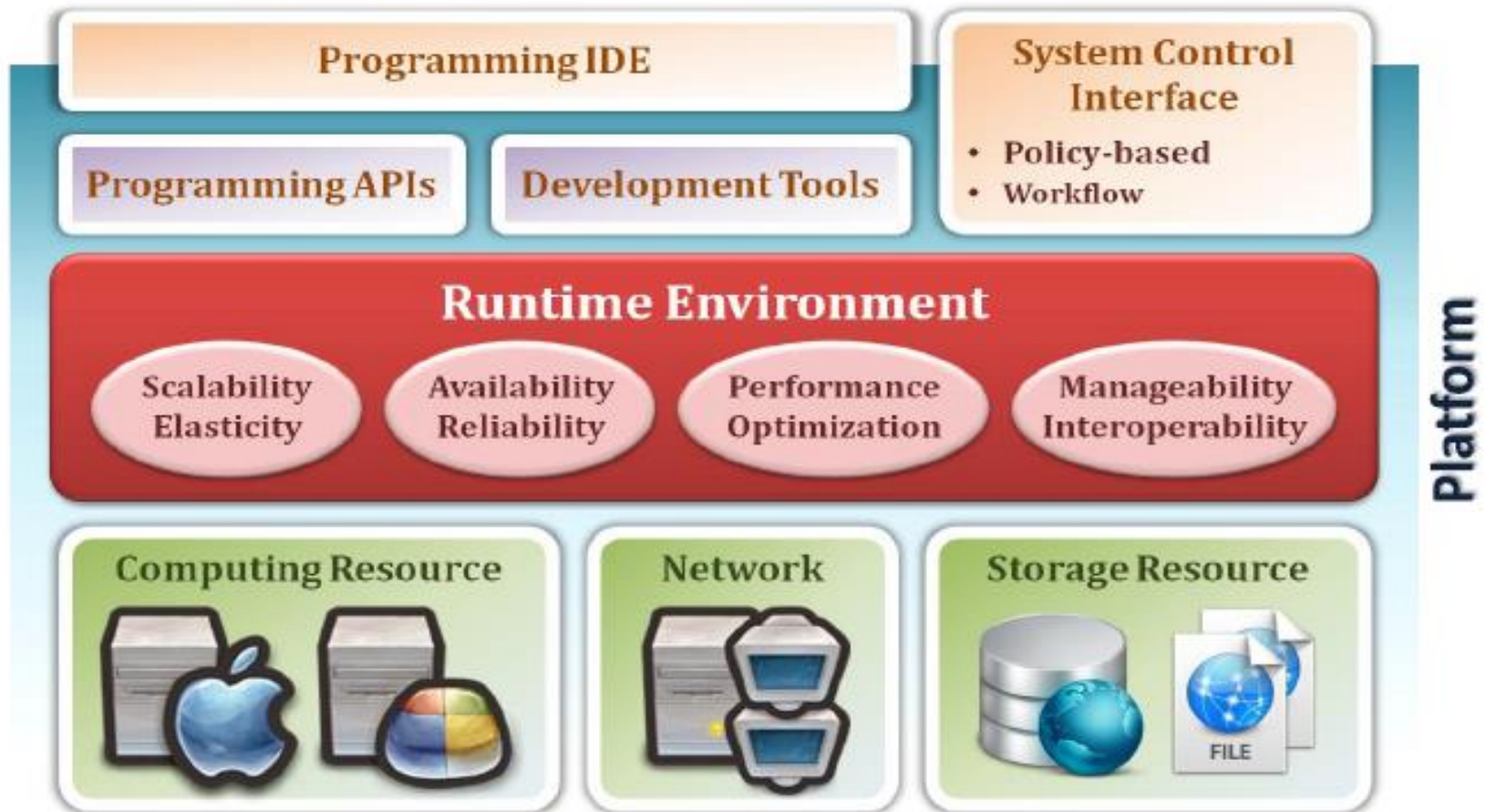
PaaS - Platform as a Service

- Designed for software developers who develop applications according to the specifications of a platform, without involving factors related to hardware infrastructure
 - Example: The platform is one that dynamically allocates resources if the application is widely used
- The consumer does not have access to the Management of cloud infrastructure (network, server, operating system or storage), but has control over the applications developed and on some configurations related to application hosting
- PaaS provides a standardized interface and a platform for SaaS level
- Example:
 - Google AppEngine – the applications run using Google infrastructure
 - Microsoft Windows Azure – PaaS level
 -

Cloud Computing | Services

PaaS - Platform as a Service

- Runtime Environment: collection of software services



Cloud Computing | Services

PaaS - Platform as a Service

- **Programming IDE**

- IDE includes all the functionality that allows access to running environment and development tools, test environment and so on

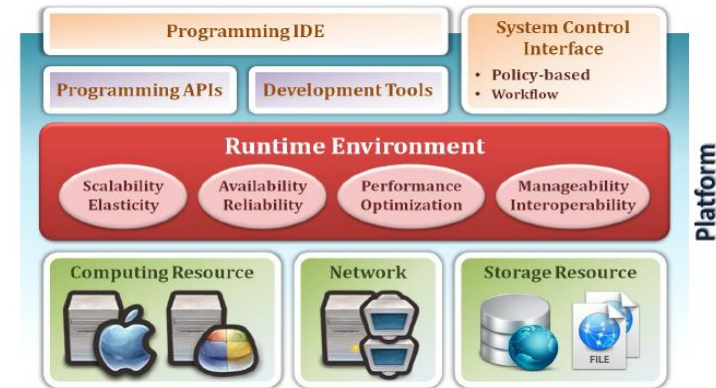
- API runtime environment varies between cloud providers, but there are common operations: computing, storage and communication.

- **System Control Interface**

- Police-Based Control: described as a principle or rule that help in making decisions for final output

- Workflow-Control:

- Describes the steps involved in resources installation and configuration
- The demon that process the workflow offer in an efficient way cloud resources



Cloud Computing | Services

PaaS - Platform as a Service

Runtime Environment supports properties as:

performance and optimization

- *scalability and elasticity*
- *availability and reliability*
- *manageability and interoperability*



Cloud Computing | Services

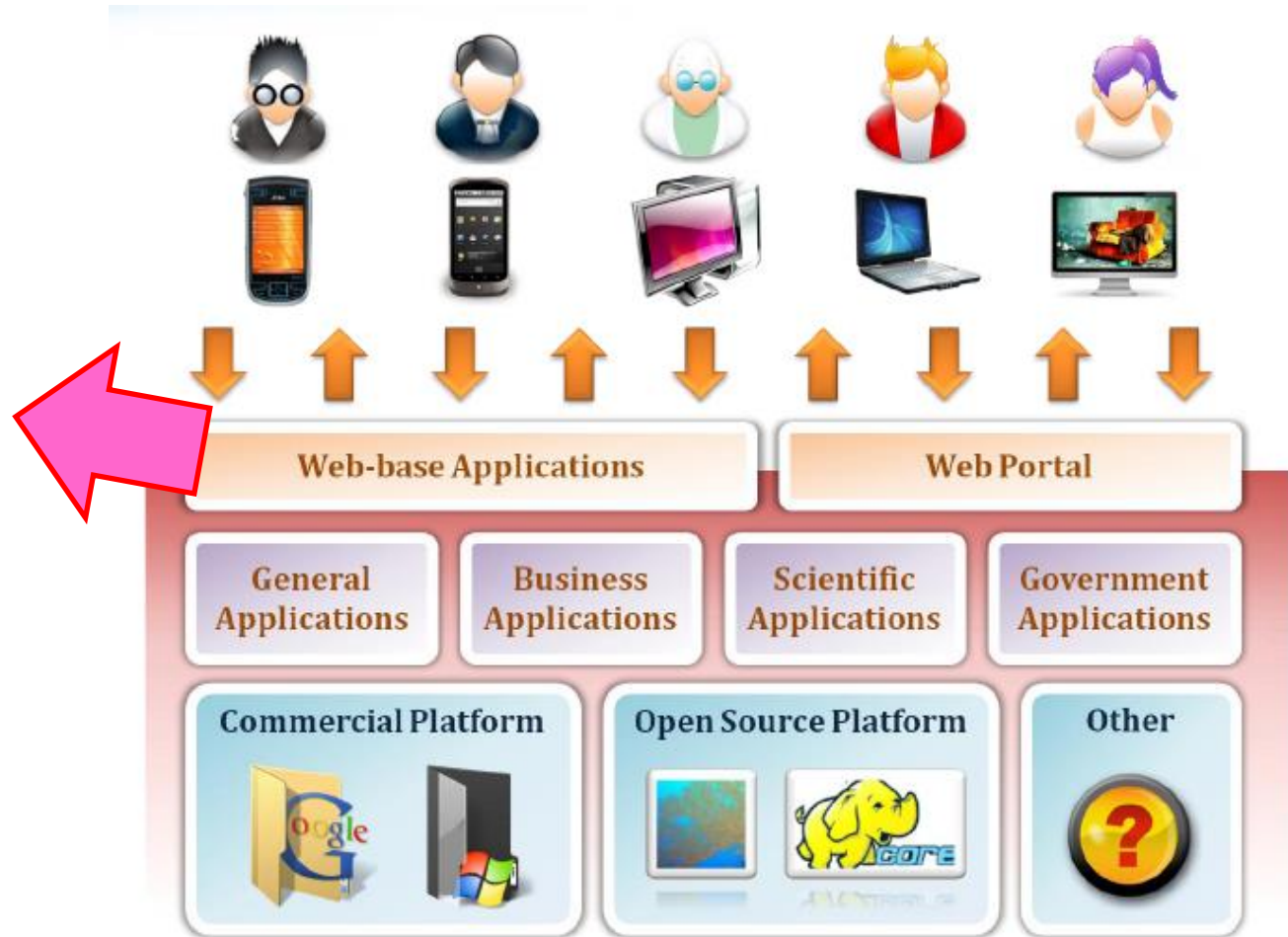
SaaS - Software as a Service

- The highest visible level from the Cloud for end-users; it provides software applications exposed as Web interfaces or Web services;
- Applications are available on a wide variety of thin clients
- “SaaS is software that is owned, delivered and managed remotely by one or more providers and that is offered in a pay-per-use manner” (Mertz 2007)
- Usually SaaS users do not know the details of the infrastructure
- Examples:
 - Google Apps (Google Mail, Google Drive, Google Spreadsheets,),
 - Live Mesh (Microsoft),....
 - Salesforce.com
 - EyeOS (<http://www.eyeos.com/>)
 -

Cloud Computing | Services

SaaS - Software as a Service

Assured proprieties:
accessibility and
portability



Cloud Computing

Deployment models

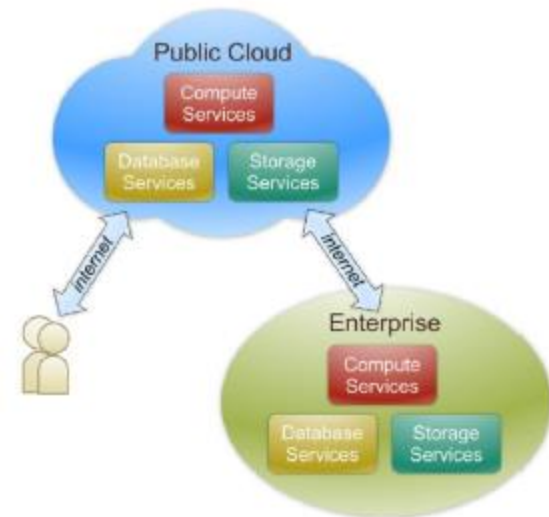
- In terms of data center owner
 - Public Cloud
 - Private Cloud
- In terms of how many cloud environments are integrated (*multiple-Cloud environments*)
 - *Community Cloud*
 - *Hybrid Cloud*

Cloud Computing

Deployment models

Public Cloud

- “A Public Cloud is data center hardware and software run by third parties, e.g. Google and Amazon, which expose their services to companies and consumers via the Internet” (IBM, 2009)
- It is available in pay-per-use manner
- Known as the external cloud and multi-tenant cloud
- Features
 - Homogeneous infrastructure
 - General rules
 - Shared resources and multi-tenancy
 - Rented infrastructure
 - It involves operating expenses

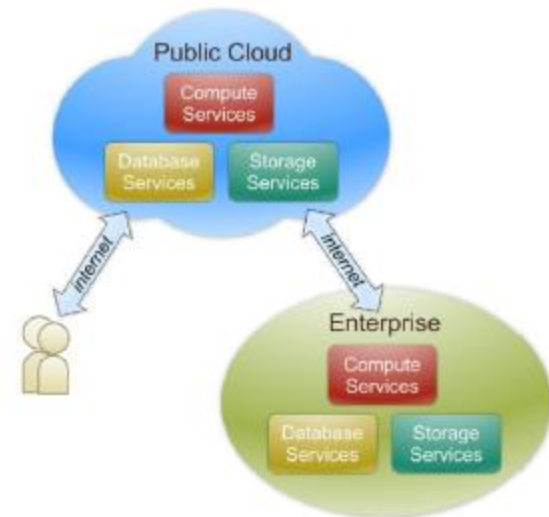


Cloud Computing

Deployment models

Private Cloud

- Motivation: limiting the risks associated with a Public Cloud
- Cloud infrastructure is used for a single organization; the cloud management can be done by another organization
- Also called internal cloud or on-premise cloud, is based on virtualization of the existing infrastructure from the organization => more efficient use of resources
- Features
 - Heterogeneous infrastructure
 - Personalized and tailored policy
 - Dedicated resources
 - Infrastructure in-house
 - Involve capital expenditure



Cloud Computing

Deployment models

Community Cloud

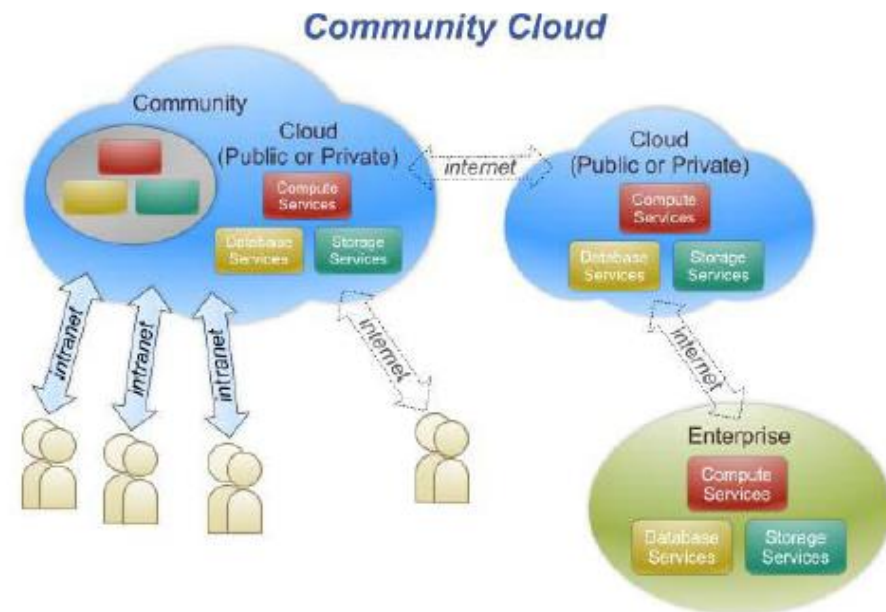
- Multiple organizations sharing their cloud infrastructure to achieve a common goal
- Also called Federation of Clouds, each cloud is independent, but can interoperate (exchange data and computing resources) with other clouds through standard interfaces
- Example: RESERVOIR
- Standardization- > Open Cloud

Computing Interface Working Group

(<http://occi-wg.org/about/specification/>)

- Open Grid Forum (OGF);

DMTF - <http://dmtf.org/standards/cloud>

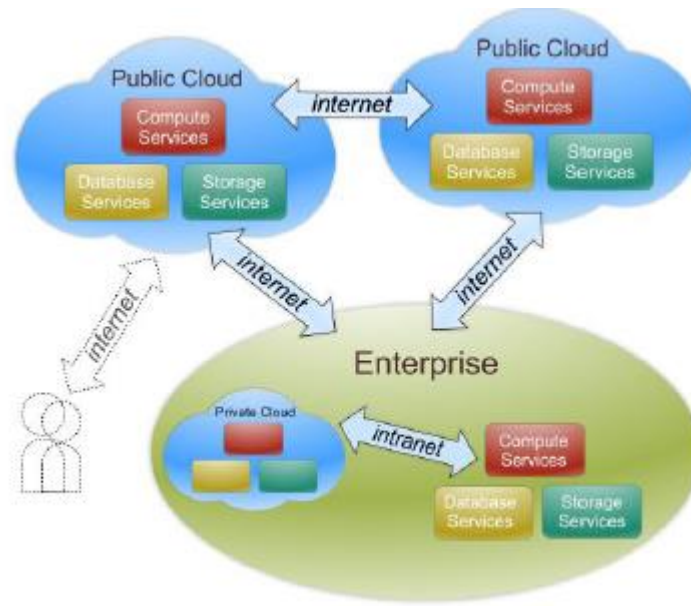


Cloud Computing

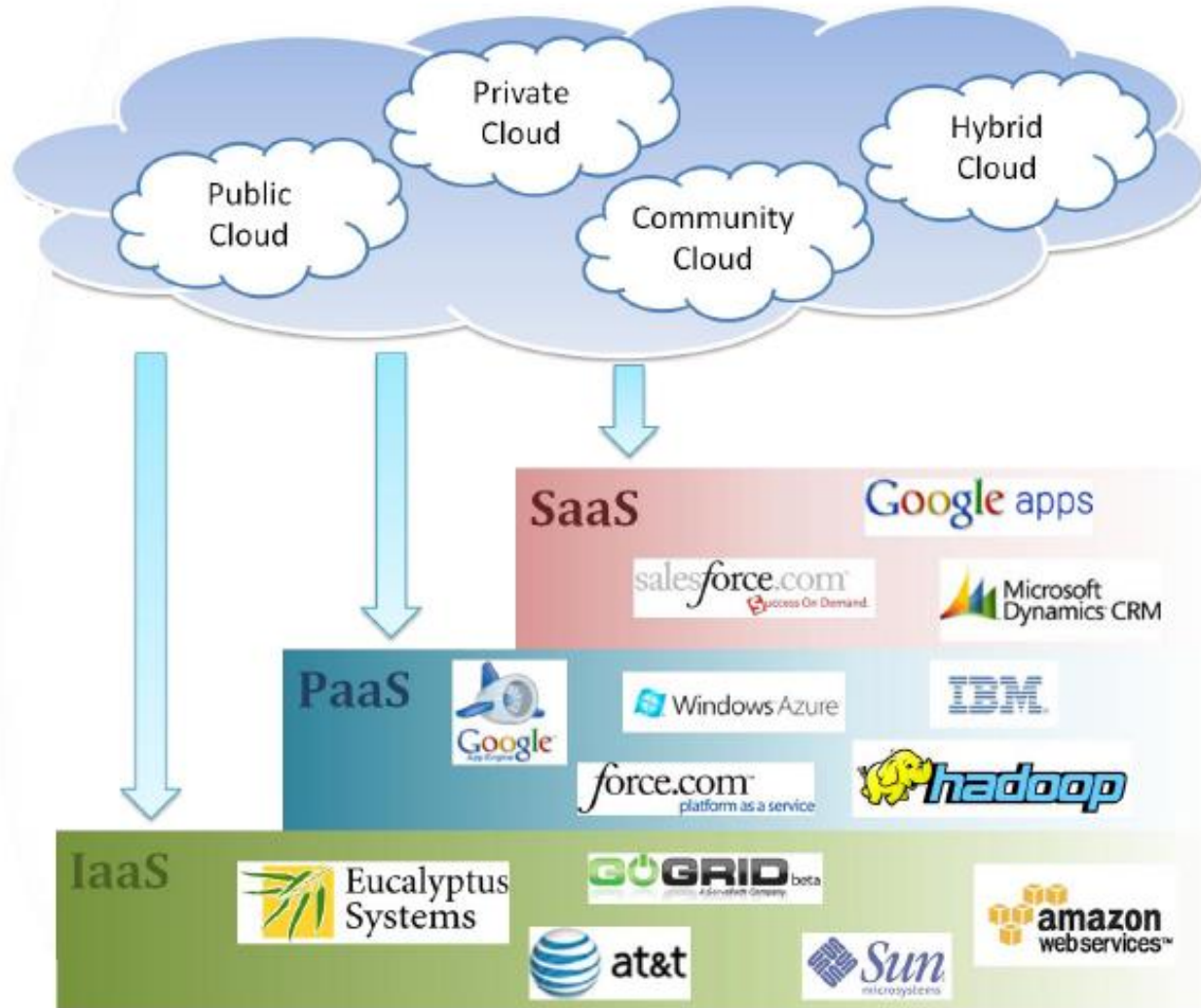
Deployment models

Hybrid Cloud

- The infrastructure consists of multiple clouds (private, community, public) that remain unique entities , but they are linked together with standardized or proprietary technologies that ensure data and application portability (e.g. in situations where load balancing is required)



Cloud Ecosystem



Bibliography

- Massimo Cafaro, Givani Aloisio, Grids, Clouds and Virtualization, 2011
- Katarina Stanoevska Slabeva, Thomas Wozniak, Grid and Cloud Computing - A Business Perspective on Technology and Applications, 2010, Editors Santi Ristol, Springer-Verlag Berlin Heidelberg
- Open Cloud Computing Interface - <http://occi-wg.org/>
- RESERVOIR - <http://ercim-news.ercim.eu/en83/special/reservoir-a-european-cloud-computing-project>
- DMTF - <http://dmtf.org/standards/cloud>
- LIBVRT - <http://libvirt.org/apps.html>
- Chow et al., Cloud Computing: Outsourcing Computation without Outsourcing Control, 1st ACM Cloud Computing Security Workshop, November 2009
- Foster, Zhao, Raicu and Lu, Cloud Computing and Grid Computing 360-Degree Compared, 2008
- Above the Clouds: A Berkeley View of Cloud Computing, Technical Report No. UCB/EECS-2009-28, <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.htm>
- <http://my.ss.sysu.edu.cn/courses/cloud/>
- <http://blogs.idc.com/ie/?p=730>
- <http://www.slideshare.net/woorung/trend-and-future-of-cloud-computing>
- <http://ganglia.sourceforge.net/>
- <http://www.focus.com/briefs/top-10-cloud-computing-trends/>

Bibliography

- NIST (National Institute of Standards and Technology).
<http://csrc.nist.gov/groups/SNS/cloud-computing/>
- <http://aws.amazon.com/free/>
- <https://www.windowsazure.com/en-us/community/education/program/overview/>
- <http://www.ibm.com/developerworks/java/library/j-gaestorage/index.html?ca=drs->
- M. Armbrust et. al., “Above the Clouds: A Berkeley View of Cloud Computing,” Technical Report No. UCB/EECS-2009-28, University of California at Berkeley, 2009.
- R. Buyya et. al., “Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility,” Future Generation Computer Systems, 2009.
- Cloud Computing Use Cases. <http://groups.google.com/group/cloud-computing-use-cases>
- Cloud Computing Explained. <http://www.andyharjanto.com/2009/11/wanted-cloud-computing-explained-in.html>
- Multiple materiale si imagini au fost preluate de pe Internet

Summary

- Cloud Computing?
 - Definitions
 - Characteristics
- Services in cloud
- Deployment models

“Alexandru Ioan Cuza” University of Iasi
Faculty of Computer Science

Questions?

