Sistemi Operativi e Reti

Cloud Computing

Facoltà di Scienze Matematiche Fisiche e Naturali

Corso di Laurea Magistrale in Informatica

Osvaldo Gervasi

ogervasi@computer.org

Introduction

- Technologies such as clusters, grid, and now cloud computing, have all aimed at allowing access to large amounts of computing power in a fully virtualized manner, by aggregating resources and offering a single system view. In addition, an important aim of these technologies, has been delivering computing as a utility
- Utility computing describes a business model for on-demand delivery of computing power.
- Pay-as-you-go model

Introduction

- Cloud computing--> umbrella term to describe a category of sophisticated on-demand computing services.
- Computing infrastructure is seen as a "Cloud" from which business and individuals access applications from anywhere in the world, on-demand.
- The main principle of cloud is offering computing, storage and software "as a service".

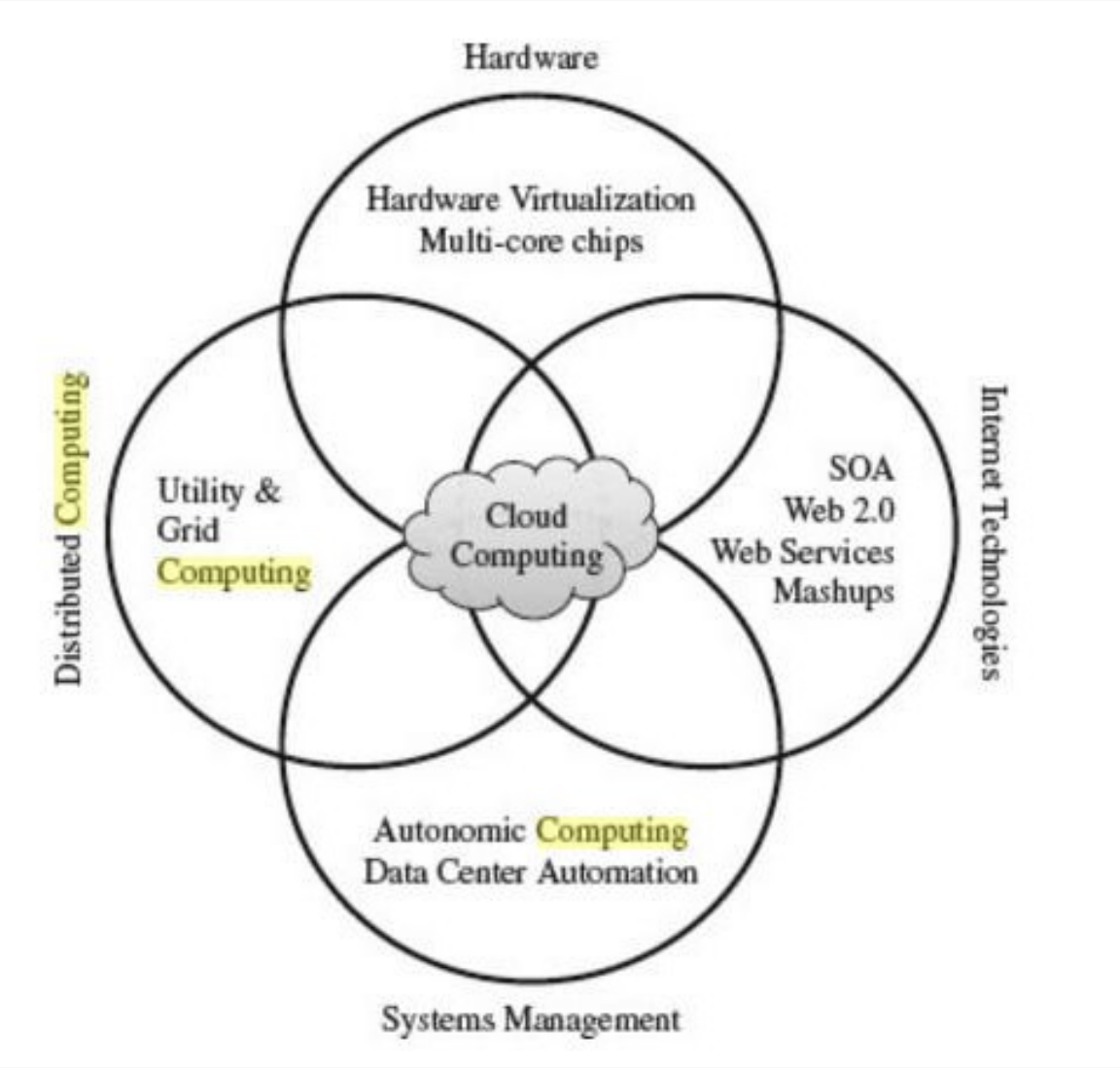
Cloud is a parallel and distributed computing system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service level agreements (SLA) established through negotiation between the service provider and the consumer. (Buyya et al)

Introduction

"The illusion of infinite computing resources;
the elimination of an up-front commitment by cloud users;
The ability to pay per use … as needed..."
(Univ. of California Berkeley)

"... a pay-per-use model for enabling available convenient,
on-demand network access to a shared pool of
configurable computing resources
(e.g. networks, servers, storage, applications, services)
That can be rapidly provisioned and released with minimal management effort or service provider interaction." (NIST)

Convergence of various advances leading to clouds



Benefits of cloud

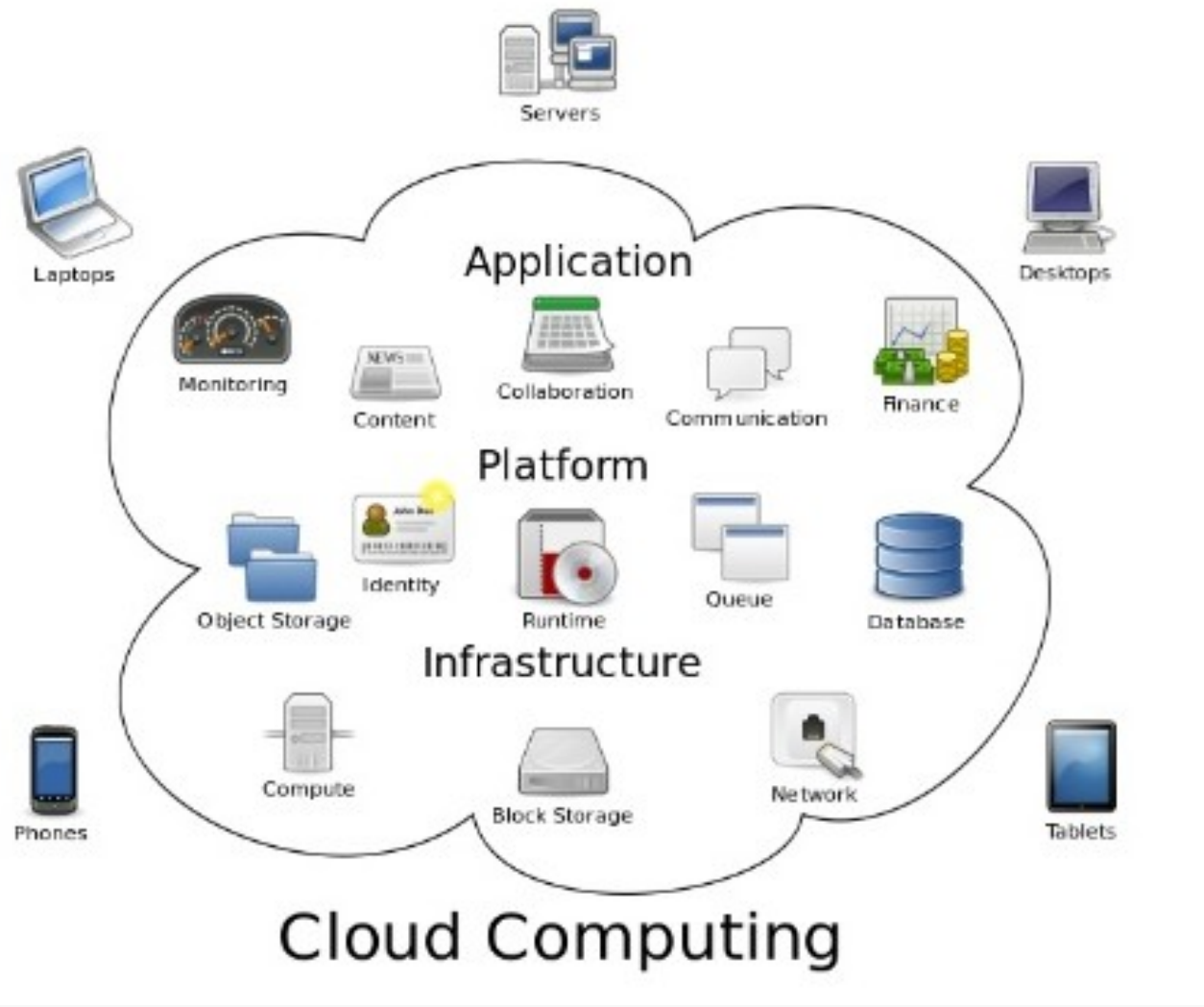
- This model brings benefits to both consumers and providers of IT services.
- Consumers can attain reduction of IT-related costs by choosing to obtain cheaper services from external providers as opposed to heavily investing on IT infrastructures and personal hiring.
- The on-demand component allows consumers to adapt their IT usage to rapidly increasing or unpredictable computing needs
- Providers achieve better operational costs, increase the efficiency, leading to a faster Return of the Investment (ROI) as well as lower Total Cost of Ownership (TCO)

SOA and Web Services

- Web Services open standards contributed significantly to the software integration
- Several technologies allow to efficiently describe, compose, and orchestrate services, publish and discover services, represent QoS parameters and ensuring security in service access.
- In a SOA software resources are packaged as services, which are well-defined, self contained modules that provide standard business functionalities and are independent of the state or context of other services.

Cloud approach

On demand delivery of infrastructure, applications and business processes in a security-rich, shared, scalable and based computing environment over the Internet for a fee”



Cloud types

Three Types?

- **IaaS** – (***Infrastructure as a Service)*** – the consumer uses "fundamental resources" such as processing power, storage, networking components or middleware. The consumer can control the operating system, storage, applications and possibly networking

- **PaaS** – **(Platform as a Service)** - the consumer uses a hosting environment for their applications and has control over the applications (and *some control over the hosting environment), but does not control the infrastructure on which they are running*

- **SaaS** – **(Software as a Service)** - the consumer uses an application, but does not control the infrastructure on which it's running (OS, hardware)

Cloud providers

Different Providers for each?

> IaaS – 

> PaaS – 

> SaaS - 

Key points

Web 2.0, Meshups

- In Web 2.0, information and services may be programmatically aggregated acting as building blocks of complex compositions, called service meshups
- Many service providers such as Amazon, Del.icio.us, Facebook and Google, make their service API publicly accessible using standard protocols such as SOAP and REST.
- In the Software as a Service (SaaS) domain cloud applications can be built as composition of other services, from the same or different providers.

EUCALYPTUS

www.eucalyptus.com

- Is an Opensource software framework for Cloud Computing, which implements the so called Infrastructure as a Service (IaaS): systems that give users the ability to run and control entire Virtual Machine instances, deployed across a variety of physical resources.
- Private Clouds enables the organizations to exploit the "elasticity" of the Cloud, maintaining the full control on the resources, monitoring their usage and planning for new services and resources.

EUCALYPTUS

- Eucalyptus enables the creation of on-premise Infrastructure as a Service clouds, with no requirements for retooling the organization's existing IT infrastructure or for introducing any specialized hardware.
- The Eucalyptus Infrastructure as a Service (IaaS) platform is fully compatible with the Amazon Web Services (AWS) API, allowing support for both on-premise and hybrid IaaS clouds.
- This compatibility allows any Eucalyptus cloud to be turned into a hybrid IaaS deployment, capable of moving workloads between AWS and on-premise data centers. Eucalyptus is compatible with a wealth of tools and applications that also adhere to the de facto AWS API standards.

Public Clouds

- Public clouds provide access to computing resources for the general public over the Internet.
- The public cloud provider allows customers to self-provision resources typically via a web service interface.
- Customer's rent access to resources as needed on a pay-as-you-go basis.
- Public clouds offer access to large pools of scalable resources on a temporary basis without the need for capital investment in data center infrastructure.

Private Clouds

- Private clouds give users immediate access to computing resources hosted within an organization's infrastructure.
- Users self-provision and scale collections of resources drawn from the private cloud, typically via web service interface, just as with a public cloud. However, because it is deployed within the organization's existing data center—and behind the organization's firewall—a private cloud is subject to the organization's physical, electronic, and procedural security measures and thus offers a higher degree of security over sensitive code and data.
- In addition, private clouds consolidate and optimize the performance of physical hardware through virtualization, and can thus markedly improve data center efficiency while reducing operational expense.

Private and Hybrid Clouds

- Eucalyptus allows to implement a "Private Cloud" or a "Hybrid Cloud". Private Clouds are based on similar technology as the Public ones (Amazon EC is the leader of such service) but all machines are behind the organization firewall.
- Hybrid Clouds combine Public and Private Cloud paradigms with a huge number of combinations.
- Private Clouds have to be open, secure, standards-based and high performance

Amazon EC2 on-demand istances

United States | Europe

Standard On-Demand Instances	Linux/UNIX Usage	Windows Usage
Small (Default)	$0.10 per hour	$0.125 per hour
Large	$0.40 per hour	$0.50 per hour
Extra Large	$0.80 per hour	$1.00 per hour
High CPU On-Demand Instances	**Linux/UNIX Usage**	**Windows Usage**
Medium	$0.20 per hour	$0.30 per hour
Extra Large	$0.80 per hour	$1.20 per hour

United States | **Europe**

Standard On-Demand Instances	Linux/UNIX Usage	Windows Usage
Small (Default)	$0.11 per hour	$0.135 per hour
Large	$0.44 per hour	$0.54 per hour
Extra Large	$0.88 per hour	$1.08 per hour
High CPU On-Demand Instances	**Linux/UNIX Usage**	**Windows Usage**
Medium	$0.22 per hour	$0.32 per hour
Extra Large	$0.88 per hour	$1.28 per hour

Amazon EC2 reserved istances

United States | Europe

Linux/UNIX	One-time Fee		
Standard Reserved Instances	**1 yr Term**	**3 yr Term**	**Usage**
Small (Default)	$325	$500	$0.03 per hour
Large	$1300	$2000	$0.12 per hour
Extra Large	$2600	$4000	$0.24 per hour
High CPU Reserved Instances	**1 yr Term**	**3 yr Term**	**Usage**
Medium	$650	$1000	$0.06 per hour
Extra Large	$2600	$4000	$0.24 per hour

United States | **Europe**

Linux/UNIX	One-time Fee		
Standard Reserved Instances	**1 yr Term**	**3 yr Term**	**Usage**
Small (Default)	$325	$500	$0.04 per hour
Large	$1300	$2000	$0.16 per hour
Extra Large	$2600	$4000	$0.32 per hour
High CPU Reserved Instances	**1 yr Term**	**3 yr Term**	**Usage**
Medium	$650	$1000	$0.08 per hour
Extra Large	$2600	$4000	$0.32 per hour

Infrastructure as a Service (IaaS)

- Infrastructure as a Service (IaaS) style clouds provide access to collections of virtualized computer hardware resources, including machines, network, and storage.
- With IaaS, users assemble their own virtual cluster on which they are responsible for installing, maintaining, and executing their own software stack.

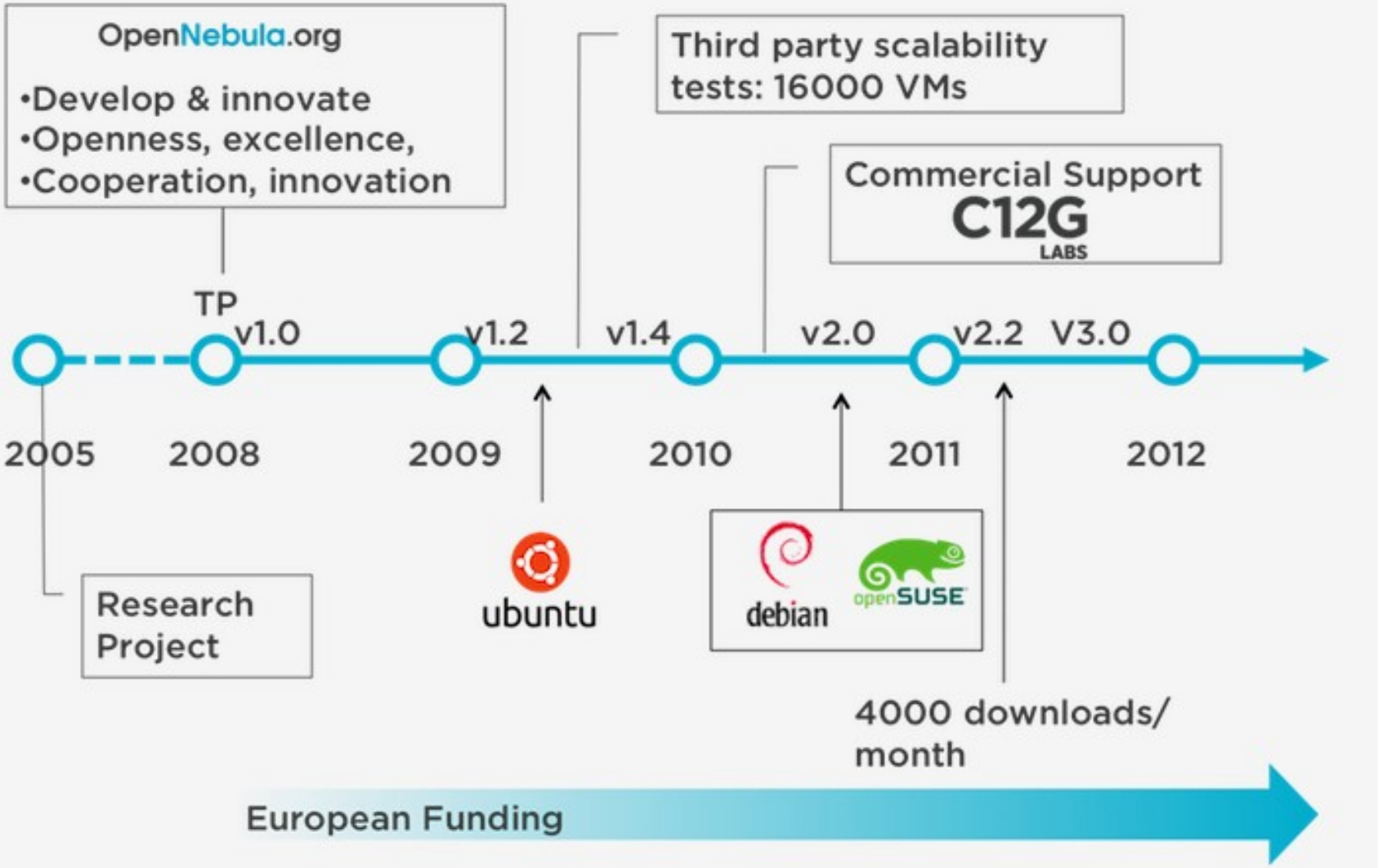
OpenNebula
opennebula.org

- OpenNebula is the industry standard open-source product **for data center virtualization**, offering the most feature-rich, customizable solution to build virtualized enterprise data centers and private cloud infrastructures on Xen, KVM and VMware deployments.
- Opennebula provides cloud consumers with choice of interfaces, from open cloud to de-facto standards, like the EC2 API.
- Opennebula is free and fully open source.

OpenNebula objectives

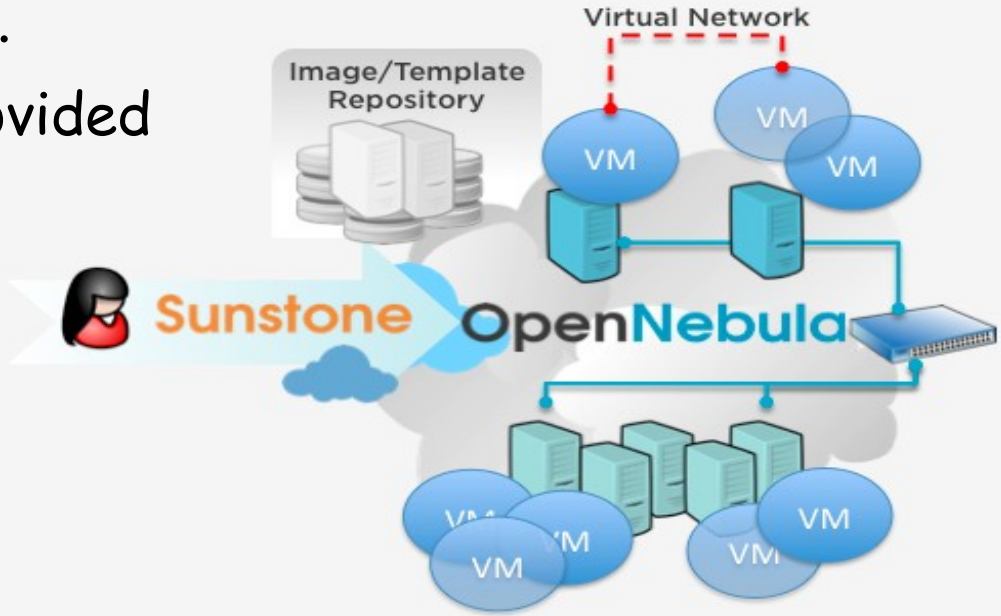
- Develop the most-advanced, highly-scalable and adaptable solution for building and managing virtualized data centers and cloud infrastructures
- Assure the stability and quality of the software distribution
- Collaborate with the most demanding users of cloud and data center management tools
- Support the ecosystem of open-source components being created around the Project
- Support the community of users and developers contributing to the Project
- Collaborate with other open-source projects and communities
- Collaborate with the main research projects in cloud computing innovation

OpenNebula history

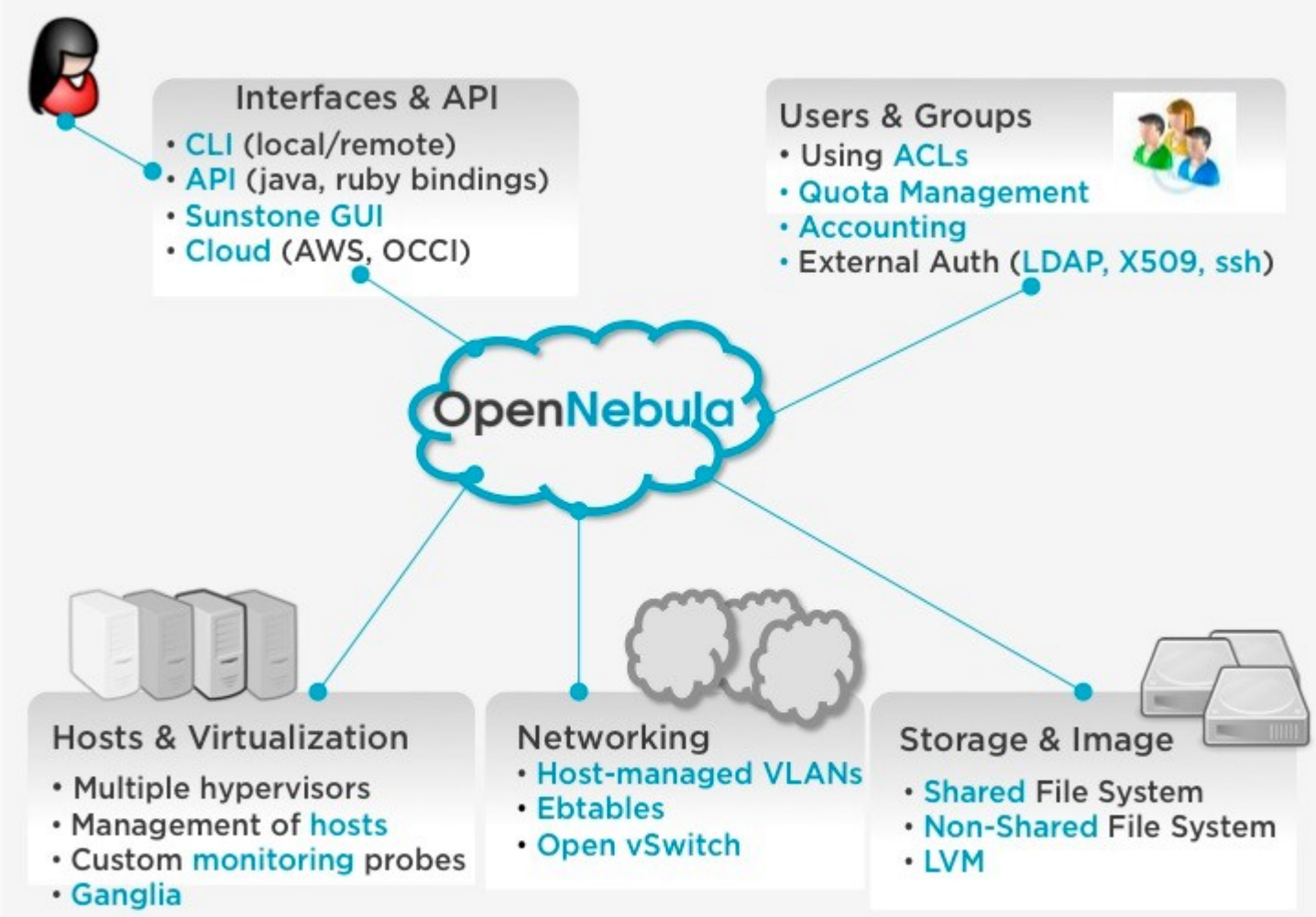


OpenNebula platform

- The **Storage system** allows to store disk images in datastores, that can be then used to define VMs or shared with other users. The images can be OS installations, or data blocks.
- The **Template Repository system** allows to register Virtual Machine definitions in the system, to be instantiated later as Virtual Machine instances.
- **Virtual Networking** is provided to interconnect Virtual Machines, they can be defined as fixed or ranged networks.



OpenNebula components



OpenNebula: advanced setups

- Multiple zones and VDCs: The OpenNebula Zones component (oZones) allows for the centralized management of multiple instances of OpenNebula, called Zones, managing in turn potentially different administrative domains.
- This zones can be further compartmentalized by grouping physical hosts in Virtual Data Centers (VDCs), so each cluster can host multiple VDCs.
- Hybrid: OpenNebula gives support to build a Hybrid Cloud is an extension of a Private Cloud to combine local resources with resources from remote Cloud providers. A whole Public Cloud provider can be encapsulated as a local resource to be able to use extra computational capacity to satisfy peak demands.

How OpenNebula operates

- OpenNebula orchestrates storage, network, virtualization, monitoring, and security technologies to enable the dynamic placement of multi-tier services (groups of interconnected virtual machines) on distributed infrastructures, combining both data center resources and remote cloud resources, according to allocation policies:
  - Management of the Network, Computing and Storage Capacity: Orchestration of storage, network and virtualization technologies to enable the dynamic placement of the multi-tier services on distributed infrastructures
  - Management of VM Life-cycle: Smooth execution of VMs by allocating the resources required for them to operate and by offering the functionality required to implement VM placement policies

How OpenNebula operates (i)

- Management of Workload Placement: Support for the definition of workload and resource-aware allocation policies such as consolidation for energy efficiency, load balancing, affinity-aware, capacity reservation…
- Management of Virtual Networks. Support for the definition of virtual networks to interconnect VMs
- Management of VM Images: Exposing of general mechanisms to transfer and clone VM images. Images can be registered before execution. When submitted, VM images are transferred to the host and swap disk images are created. After execution, VM images may be copied back to the repository

How OpenNebula operates (ii)

- Management of Information and Accounting. Provision of indicators that can be used to diagnose the correct operation of the servers and VMs and to support the implementation of the dynamic VM placement policies
- Management of Security: Definition of security policy on the users of the system, guaranteeing that the resources are used only by users with the relevant authorizations and isolation between workloads
- Management of Remote Cloud Capacity: Dynamic extension of local capacity with resources from remote providers to build hybrid or federated cloud deployments
- Management of Public Cloud Servers: Exposing most common cloud interfaces to support public cloud computing deployments

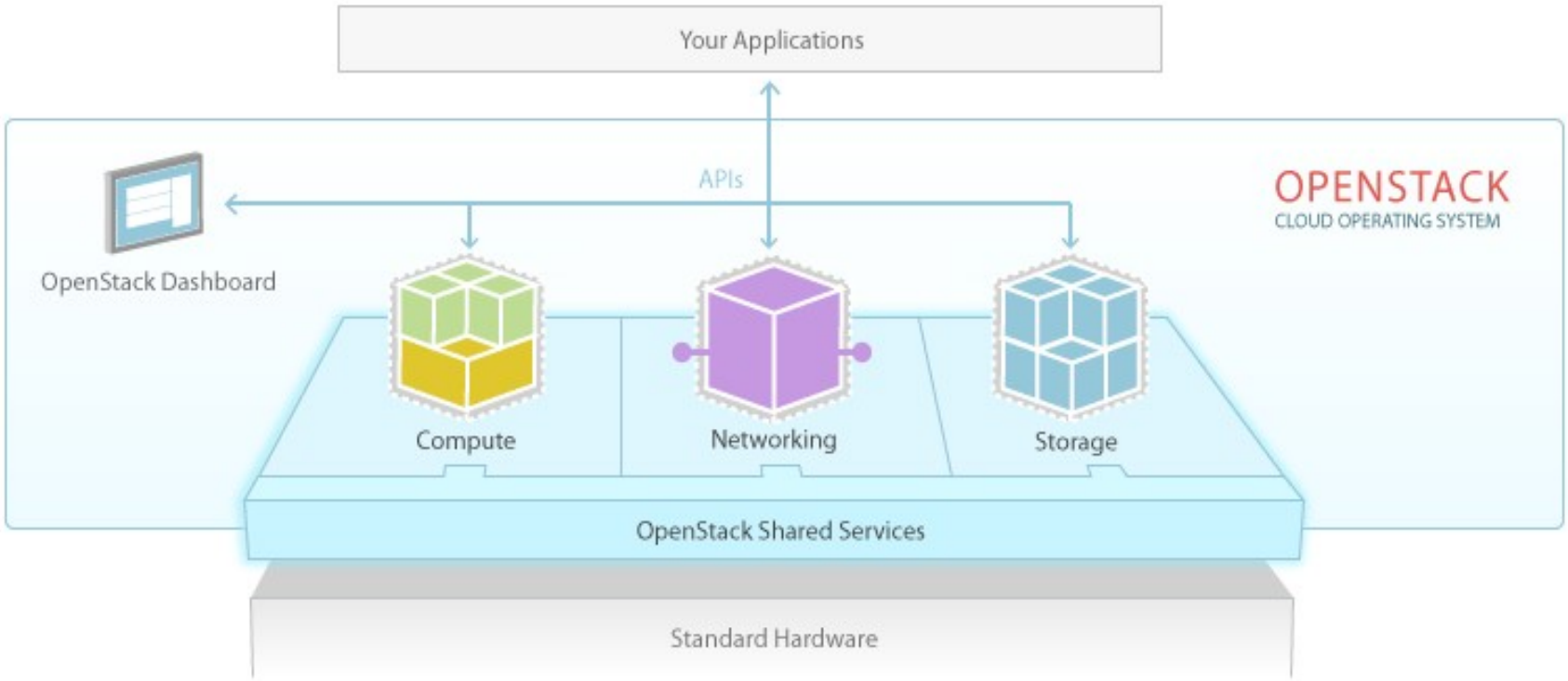
OpenStack

- OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter, all managed through a dashboard that gives administrators control while empowering their users to provision resources through a web interface.
- There are many OpenStack public clouds in production across the world, including Cloudwatt, DreamCompute, eNocloud, HP, Rackspace, and Ulticloud .

OpenStack

- The alternative to the use of public clouds (which cost money), is to install Devstack
- DevStack is a great option to install and run an OpenStack cloud on your laptop (or even inside the VM on a cloud).
- DevStack is ideal for potential users who want to see what the Dashboard looks like from an admin or user perspective, and OpenStack contributors wanting to test against a complete local environment.

OpenStack



OpenStack

- The OpenStack cloud operating system enables enterprises and service providers to offer on-demand computing resources, by provisioning and managing large networks of virtual machines.
- Compute resources are accessible via APIs for developers building cloud applications and via web interfaces for administrators and users.
- The compute architecture is designed to scale horizontally on standard hardware, enabling the cloud economics companies have come to expect.

OpenStack

- OpenStack is architected to provide flexibility as you design your cloud, with no proprietary hardware or software requirements and the ability to integrate with legacy systems and third party technologies.
- It is designed to manage and automate pools of compute resources and can work with widely available virtualization technologies, as well as bare metal and high-performance computing (HPC) configurations.

OpenStack

- Administrators often deploy OpenStack Compute using one of multiple supported hypervisors in a virtualized environment. KVM and XenServer are popular choices for hypervisor technology and recommended for most use cases.
- Linux container technology such as LXC is also supported for scenarios where users wish to minimize virtualization overhead and achieve greater efficiency and performance.
- In addition to different hypervisors, OpenStack supports ARM and alternative hardware architectures.

OpenStack

- Popular Use Cases:
  - Service providers offering an IaaS compute platform or services higher up the stack
  - IT departments acting as cloud service providers for business units and project teams
  - Processing big data with tools like Hadoop
  - Scaling compute up and down to meet demand for web resources and applications
  - High-performance computing (HPC) environments processing diverse and intensive workloads