

Cloud Platform Architecture

Acknowledgement: Prof. Rajkumar Buyya for providing figures appear in this presentation

Cloud Services on Internet



Introduction

- *Cloud computing* is a utility oriented and Internet centric way of delivering IT services on demand.
- These services cover the entire computing stack: from the hardware infrastructure packaged as a set of virtual machines to software services such as development platforms and distributed applications.

Introduction

*“**Cloud computing** is a computing paradigm shift where computing is moved away from personal computers or an individual application server to a “cloud” of computers. Users of the cloud only need to be concerned with the computing service being asked for, as the underlying details of how it is achieved are hidden. This method of distributed computing is done through pooling all computer resources together and being managed by software rather than a human.”*

Source: http://dbpedia.org/page/Cloud_computing

Definition of Cloud

US National Institute of Standards and Technology

Cloud computing is a model for enabling convenient, on demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction

Early Definitions of Cloud

- *Powerful **services and applications** are being integrated and **packaged** on the web in what the industry now calls 'cloud computing'*
- *Providing services on **virtual machines** allocated on top of a large physical machine pool*

Early Definitions of Cloud

- *A Cloud is a type of parallel and distributed system consisting of **interconnected and virtualized** computers that are dynamically provisioned and presented as one or more unified computing resources based on Service Level Agreement (SLA) established through **negotiation** between the service provider and consumers (2009)*

Interpretation of Cloud

- Clouds appears to be combination of Cluster and Grids, but it is not the case. Clouds are *next generation compute/data centers with node are virtualized using hypervision technology*, and allocates resources (Infrastructure/Software/Platforms) with business model to the user dynamically to meet specific objective based on the SLA.
- SLA is established using standard web services technology such as SOAP (WSLA) or REST(**RE**presentational **St**ate **T**ransfer).

Characteristics of Cloud Computing

- **Virtual** – Physical location and underlying infrastructure details are transparent to users
- **Scalable** – Able to break complex workloads into pieces to be served across an incrementally expandable infrastructure
- **Efficient** – Services Oriented Architecture for dynamic provisioning of shared compute resources
- **Flexible** – Can serve a variety of workload types – both consumer and commercial
- **Elastic** – Can shrink and grow based on the requirements: follow the principle of create, use and destroy principle

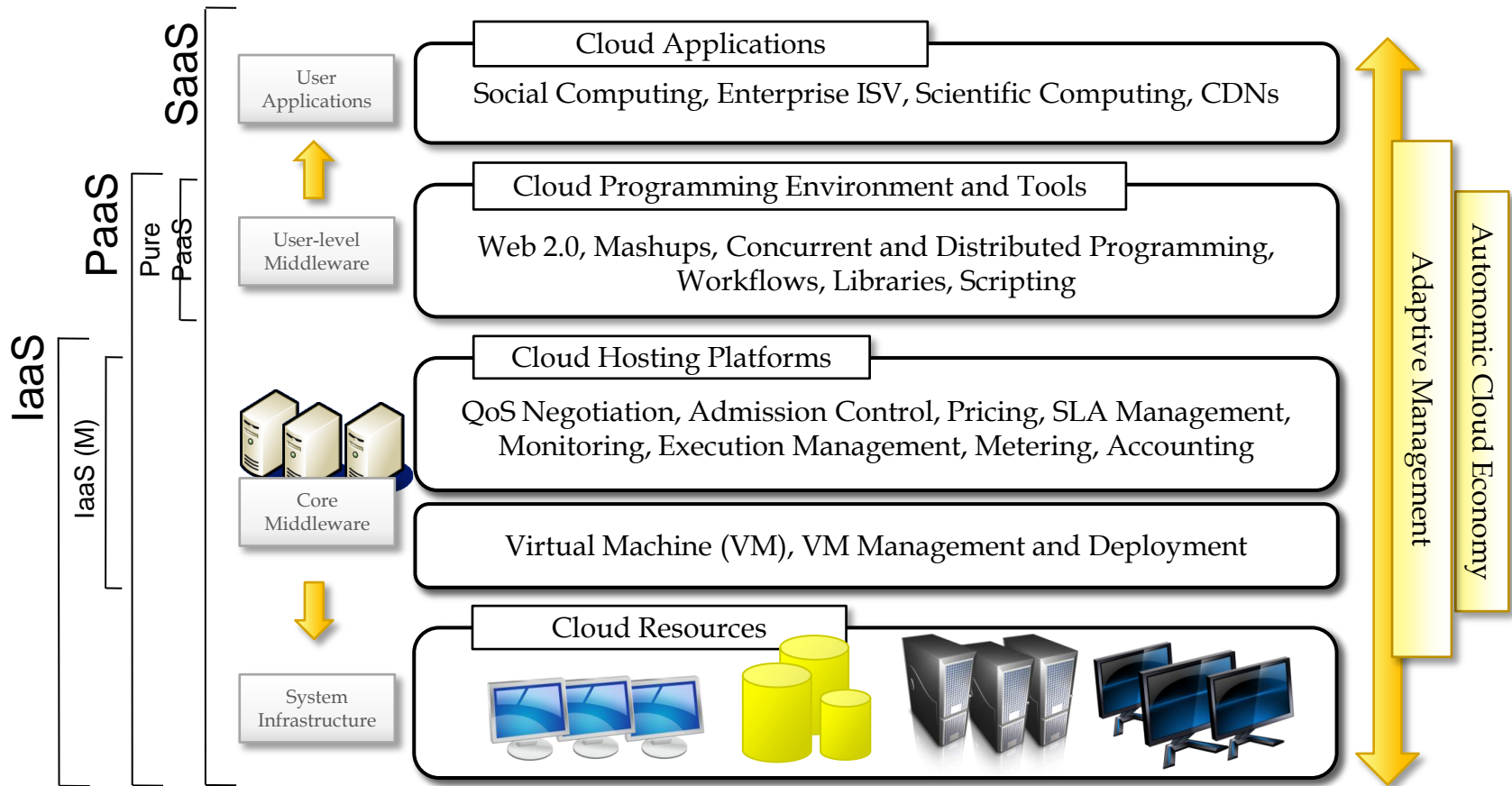
Some benefits of Clouds to Users

- **No Infrastructure investments**
 - No need to buy HW, Network Devices, Power
- **On demand service**
 - Get when and what you need
- **Proper price**
 - Only pay what and how much you use
- **Efficient resource allocation**
 - Different people live in different time zones
- **High availability**
 - One can use it any time

Cloud Reference Model

- Cloud computing supports any IT service that can be consumed as a utility and it is delivered through the network, most likely the Internet.
- Such characterization includes quite different aspects: infrastructure, development platforms, application and services.

Cloud Computing Architecture



Cloud Services

Cloud computing is Internet-based computing, whereby services like Software, Platform, Infrastructure are provided to clients on demand.

- **Infrastructure as a Service (IaaS, some call it HaaS)**
 - CPU cycles, Storage **Ex. Amazon.com**
- **Platform as a Service (PaaS)**
 - GoogleApp Engine, Azure Microsoft, Aneka
- **Software as a Service (SaaS)**
 - SalesForce.Com, Office 365

Infrastructure as a Service (IaaS)

- The capability provided to the consumer is to rent processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications.
- The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly select networking components (e.g., firewalls, load balancers).

Platform as a Service (PaaS)

- **Platform as a Service (PaaS)** model enables the user to deploy user-built applications onto a virtualized cloud platform.
- PaaS includes middleware, databases, development tools, and some runtime support such as Web 2.0 and Java.

Platform as a Service (PaaS)

- The platform includes both hardware and software integrated with specific programming interfaces.
- The provider supplies the API and software tools (e.g., Java, Python, Web 2.0, .NET).
- The user is freed from managing the cloud infrastructure.

Software as a Service (SaaS)

- **Software as a Service (SaaS)** is based on a simple model. It allows the delivery of an application via a **subscription model** over the Internet.
- In the SaaS model, the customer does not take ownership of the software but rather **'subscribes'** to a comprehensive solution that is delivered remotely over the Internet. No infrastructure investments, no maintenance costs.

IaaS, PaaS, SaaS

Category	Characteristics	Product Type	Vendors & Products
<i>SaaS</i>	Customers are provided with applications that are accessible anytime and from anywhere.	Web applications and services (Web 2.0)	SalesForce.com (CRM); Clarizen.com (Project Management); Google Apps...
<i>PaaS</i>	Customers are provided with a platform for developing applications hosted in the Cloud.	Programming APIs and frameworks; Deployment Systems.	Google AppEngine; Microsoft Azure; Manjrasoft Aneka; Data Synapse...
<i>IaaS/HaaS</i>	Customers are provided with virtualized hardware and storage on top of which they can build their infrastructure.	Virtual machines management infrastructure; Storage management; Network management.	Amazon EC2 and S3; GoGrid; Nirvanix...

SaaS:

- Gov-Apps, Internet Services
- Blogging/Surveys/Twitter, Social Networking
- Information/Knowledge Sharing (Wiki)
- Communication (e-mail), Collaboration (e-meeting)
- Productivity Tools (office)
- Enterprise Resource Planning (ERP)

PaaS:

- Application Development, Data, Workflow, etc.
- Security Services (Single Sign-On, Authentication, etc.)
- Database Management
- Directory Services

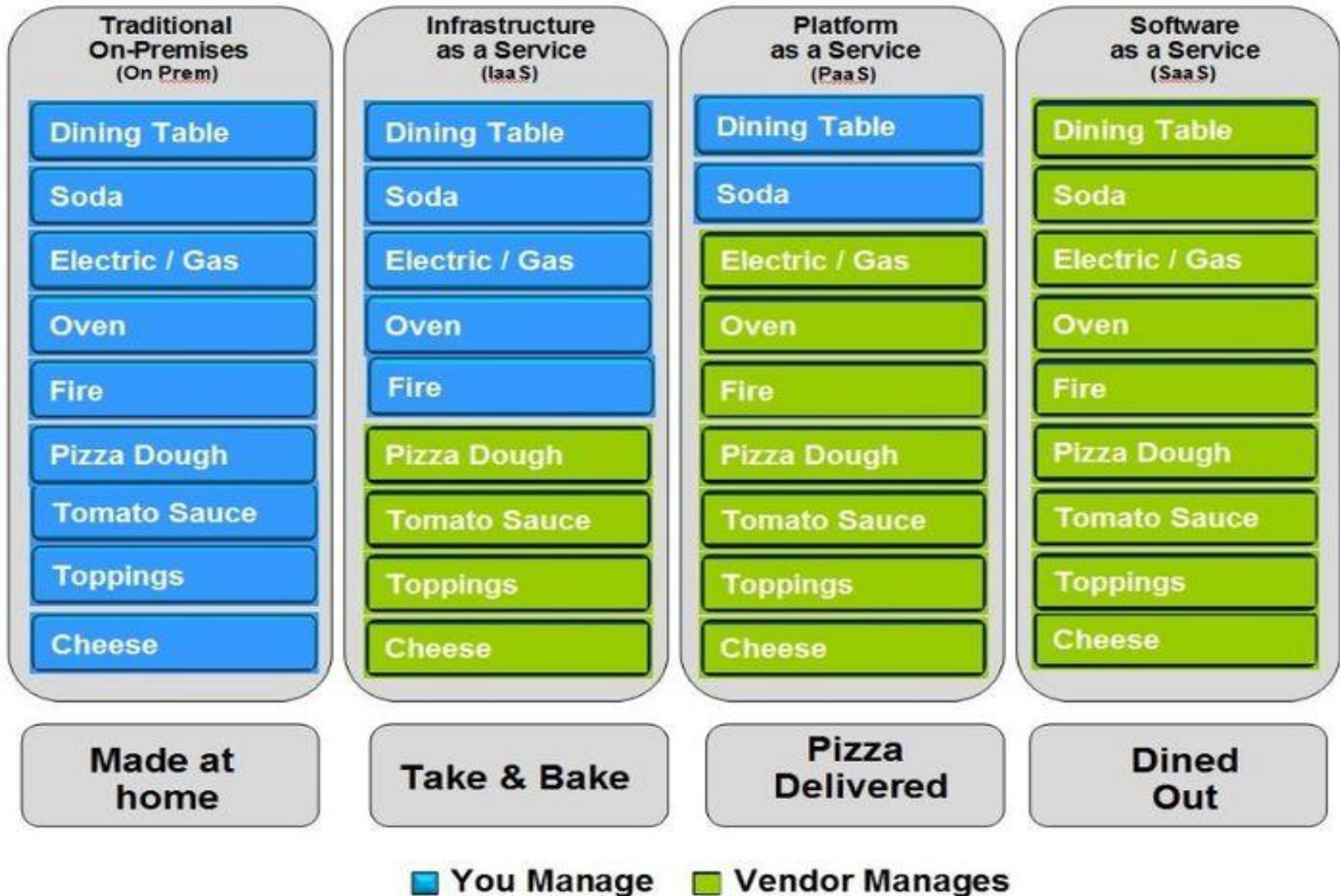
IaaS:

- Networks, Security, Mainframes, Servers, Storage
- Telecom Carrier Services
- IT Facilities/Hosting Services

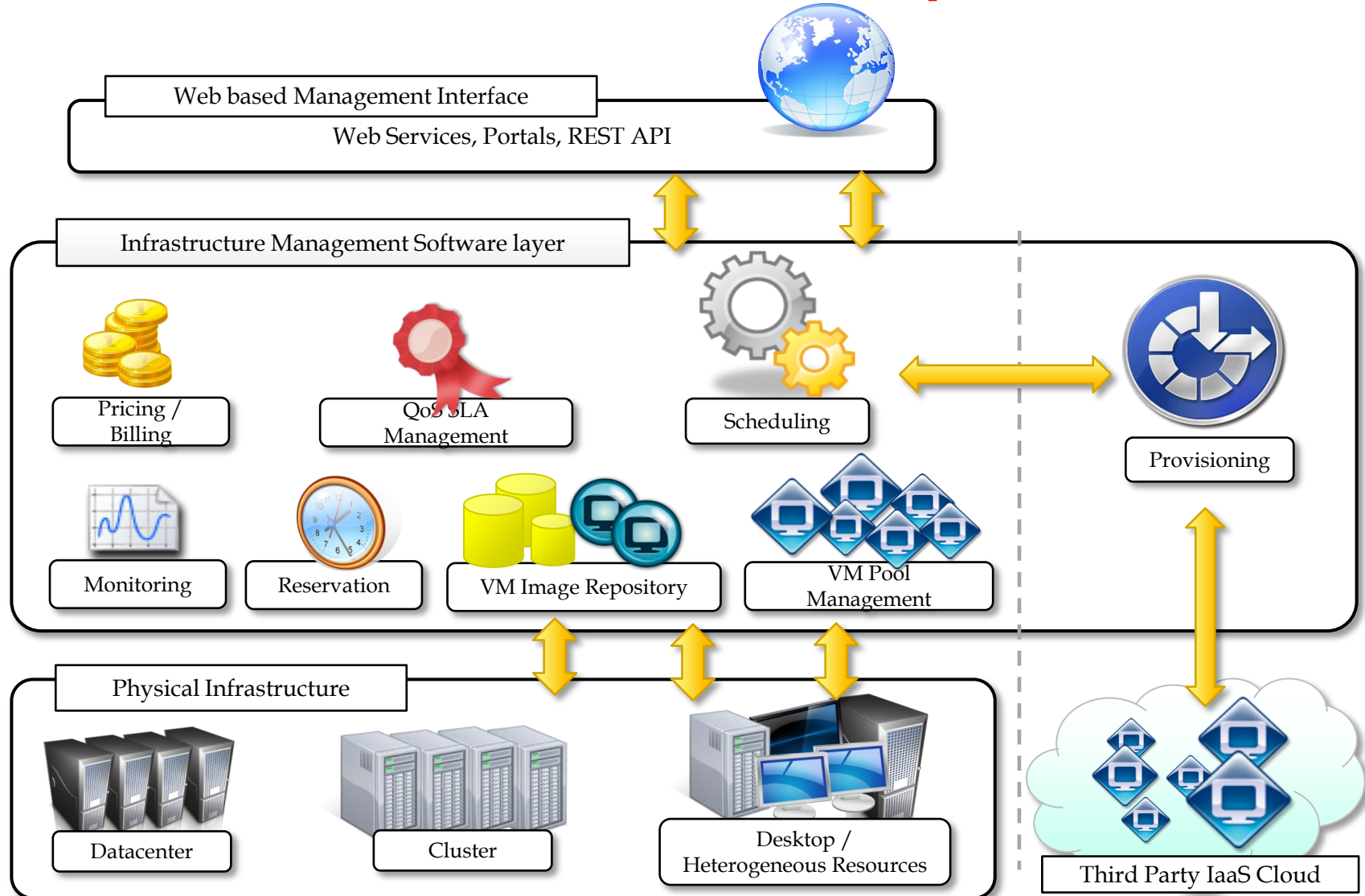
XaaS

- ***Everything as a Service*** (XaaS) is one of the most important elements of Cloud computing:
- Cloud services from different providers can be composed together in order to provide a completely integrated solution covering all the computing stack of a system.

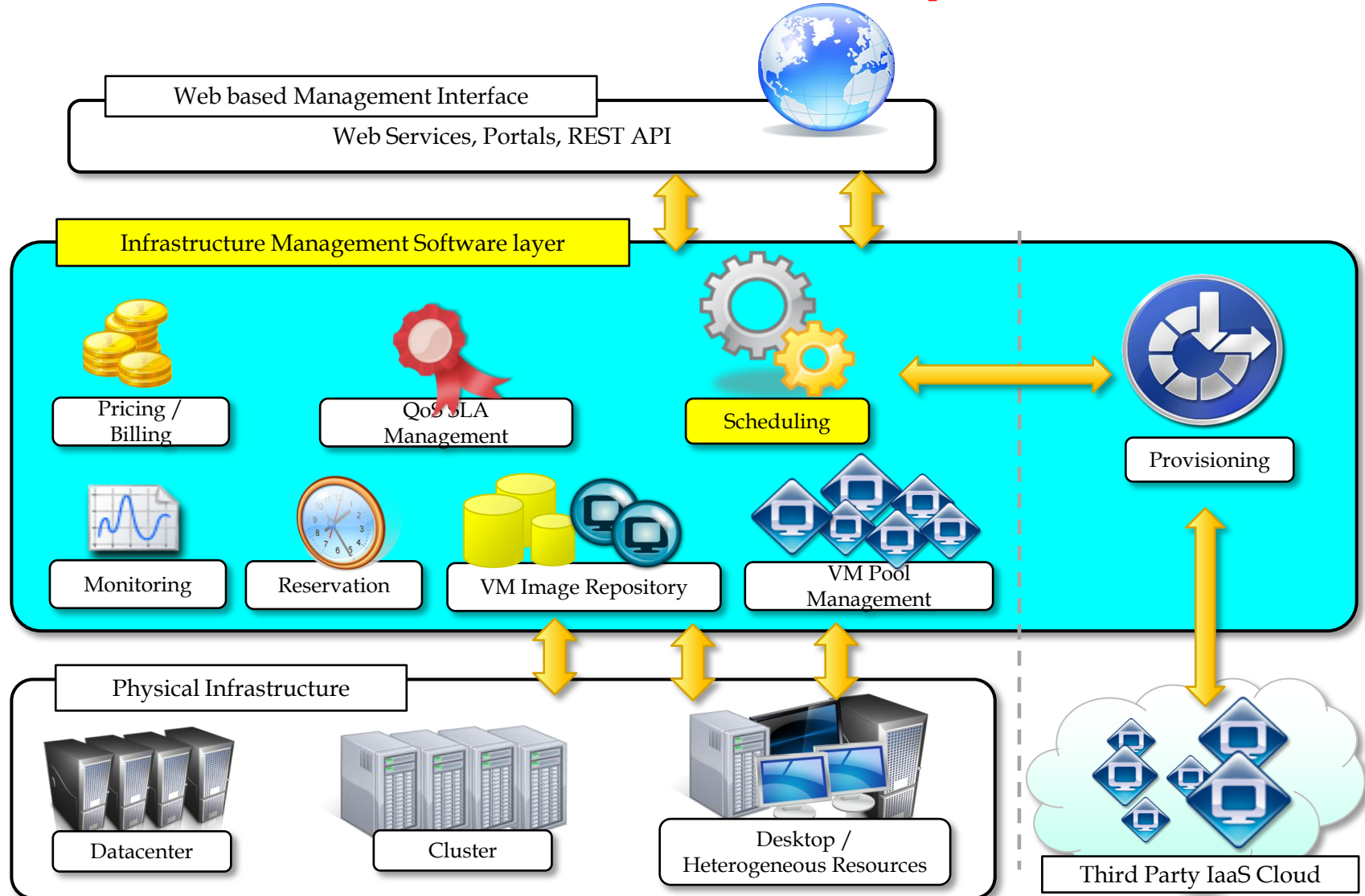
Pizza as a Service



Infrastructure as a Service implementation



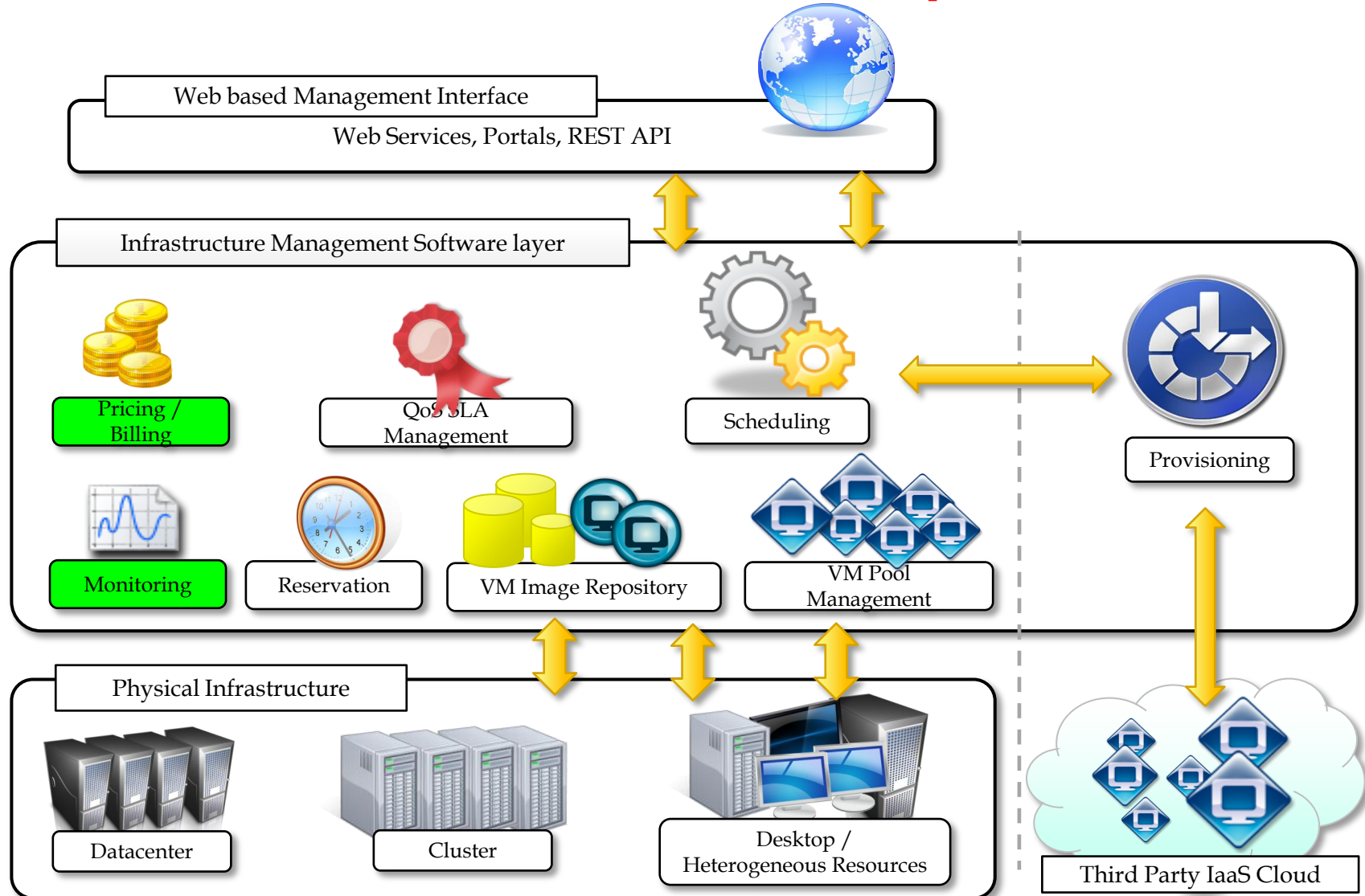
Infrastructure as a Service implementation



IaaS Conti...

- The core features of an Infrastructure-as-a-Service solution are implemented in the *infrastructure management software layer*.
- The management of the VMs is the most important function performed by this layer.
- A central role is played by the scheduler, which is in-charge of allocating the execution of virtual machine instances.
- The scheduler interacts with the other components performing different tasks:

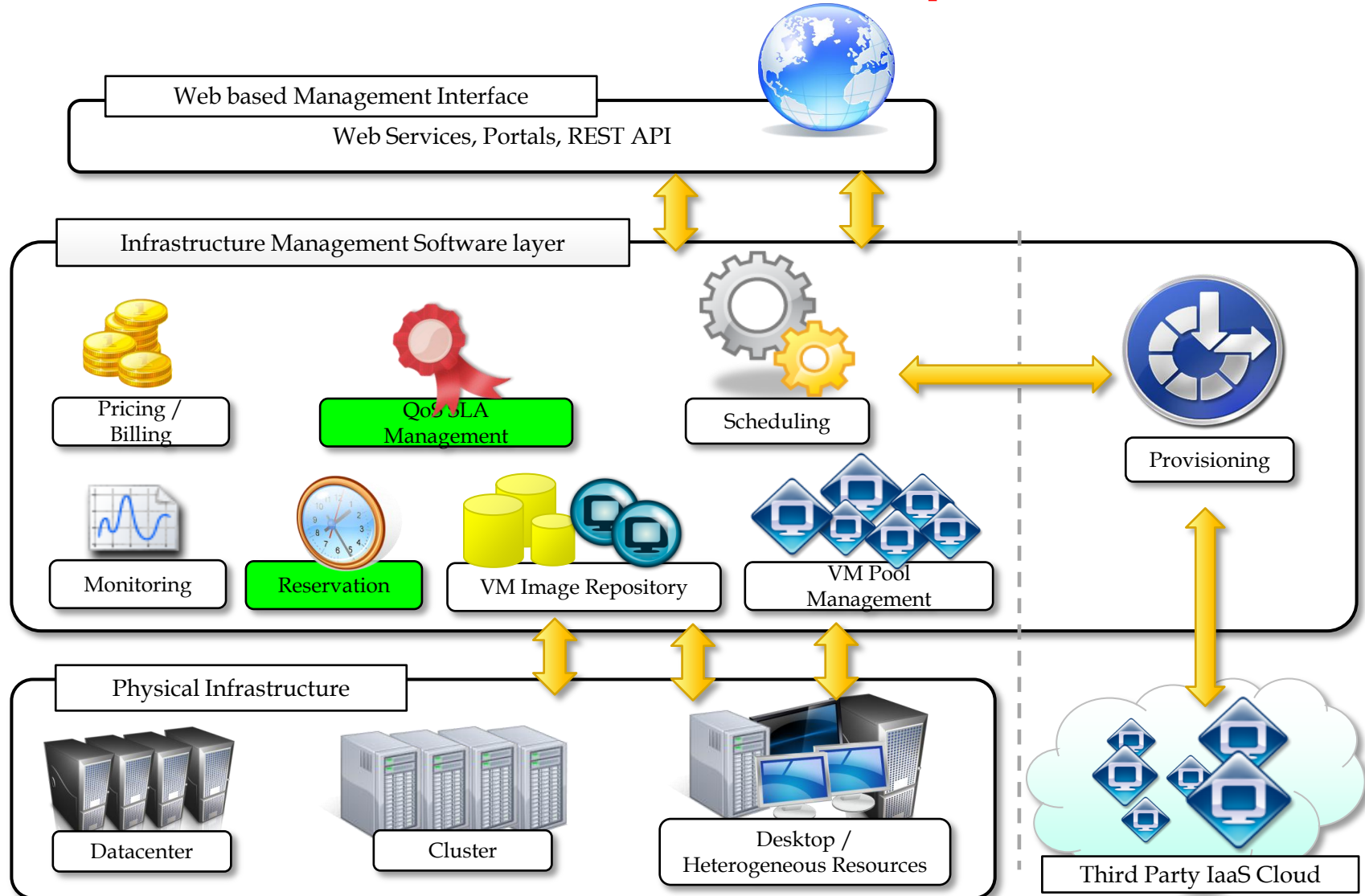
Infrastructure as a Service implementation



IaaS Conti...

- The *pricing / billing* component takes care of the *cost of executing* each virtual machine instance and maintains data that will be used to charge the user.
- The *monitoring* component tracks the execution of each virtual machine instance and maintains data required for *reporting and analyzing* the performance of the system.

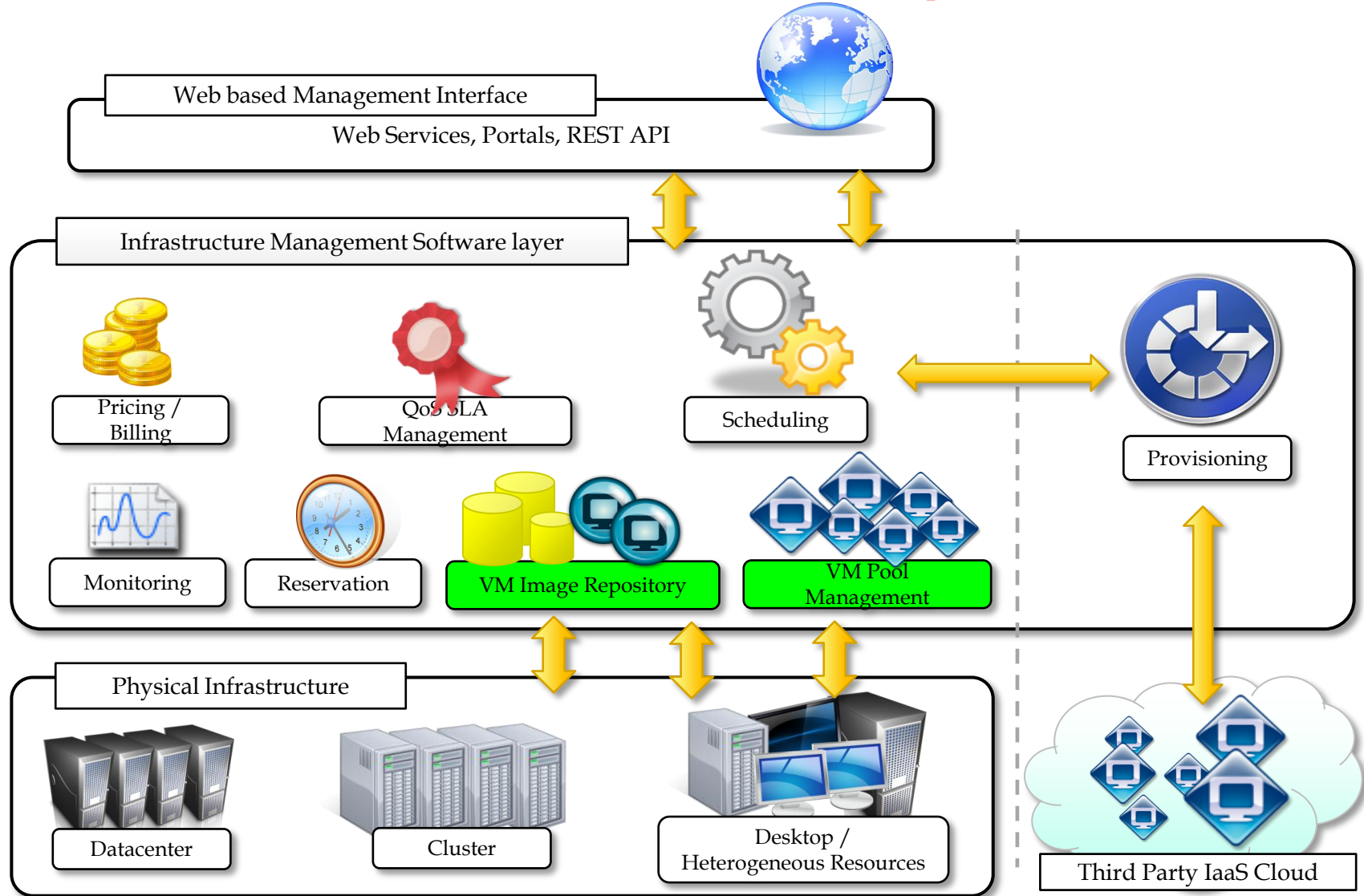
Infrastructure as a Service implementation



IaaS Conti...

- The *reservation* component stores the information of all the VM instances that **have been executed** or that **will be executed** in the future.
- In case support for QoS-based execution is provided, a *QoS/SLA management* component maintains a repository of all the service level agreements made with the users and **together with the monitoring component** is used to ensure that a given virtual machine instance is executed with the desired Quality of Service.

Infrastructure as a Service implementation



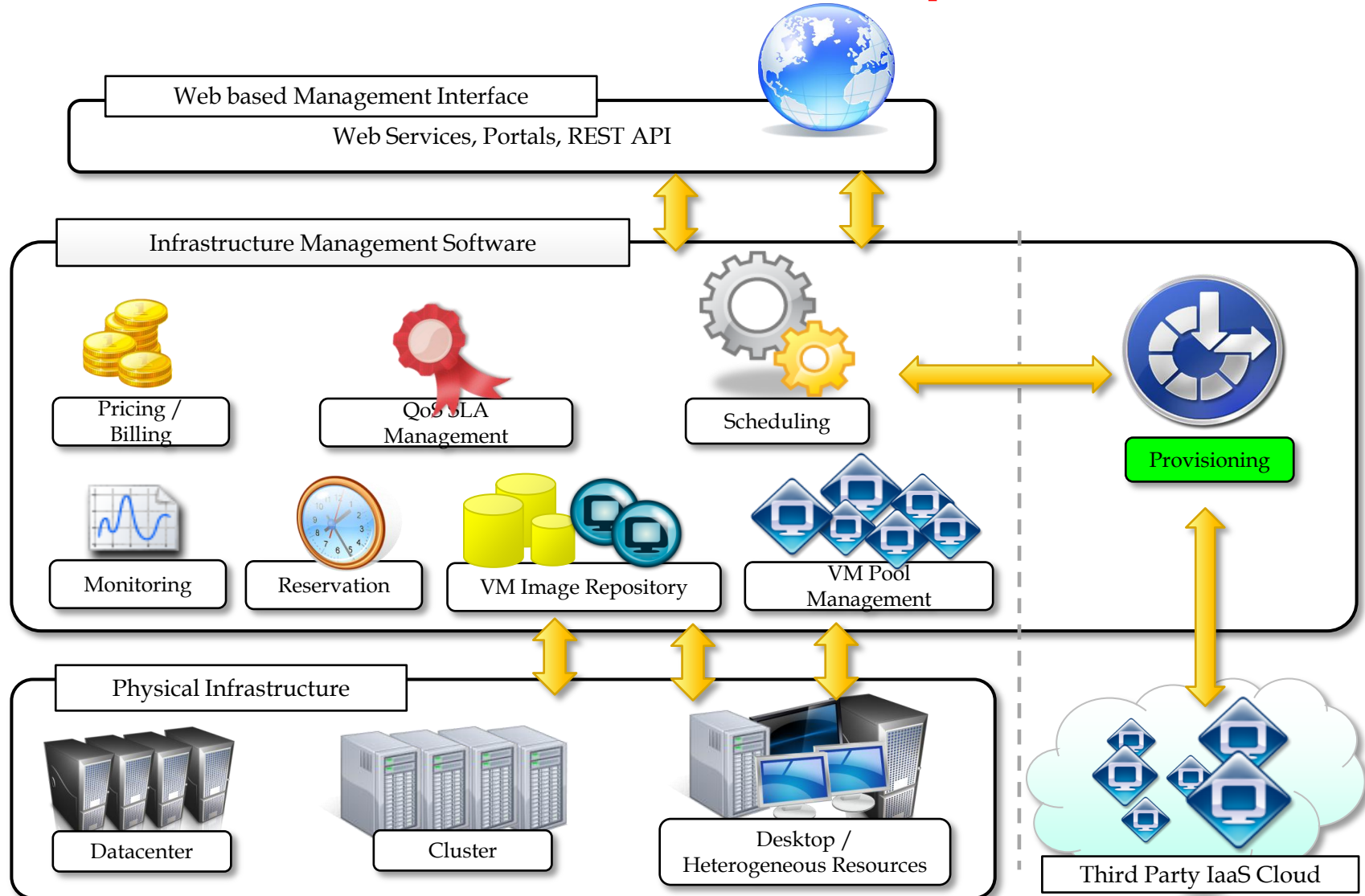
IaaS Conti...

- The *VM repository* component provides a **catalog of virtual machine images** that users can use to create virtual instances. Some implementations also allow the users to upload their specific virtual machine image.
- A *VM pool manager* component is responsible of keeping track of all the ***live instances***.

IaaS Conti...

- ***Image is static data*** containing the software (the OS and applications together with their configuration and data files etc.) that the virtual machine will run once started. It is usually stored on disk (though you can store it anywhere you like).
- ***Instance is a running virtual machine.*** It has been started from an image and it is capable of running an OS and processes, performing computations, I/O etc. Unlike static image, instance is a dynamic thing which does some work and with which you can interact.

Infrastructure as a Service implementation



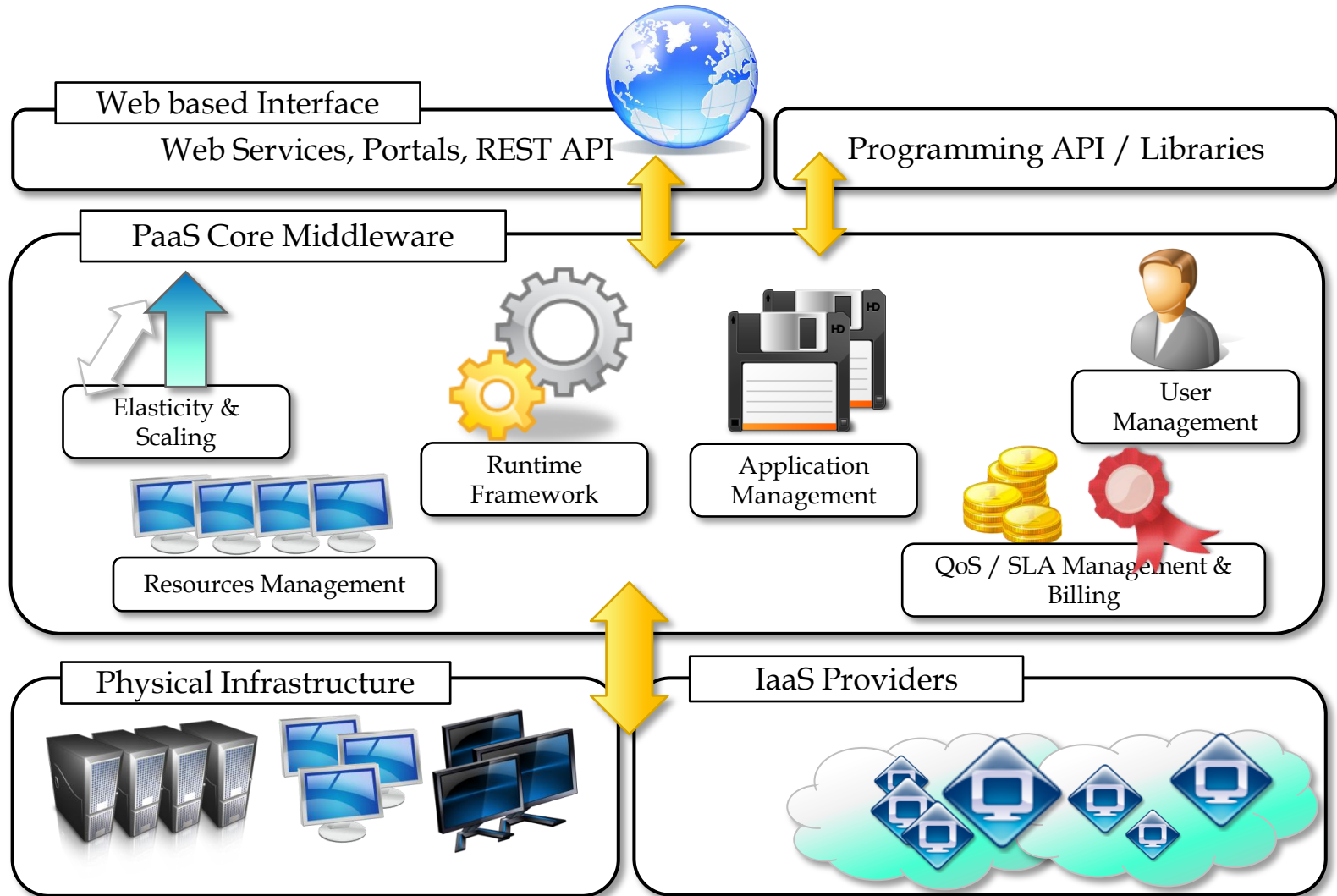
IaaS Conti...

- If the system supports the integration of additional resources belonging to a third party IaaS provider, a *provisioning* component interacts with the scheduler in order to provide a virtual machine instance that is external to the local physical infrastructure directly managed by the pool.

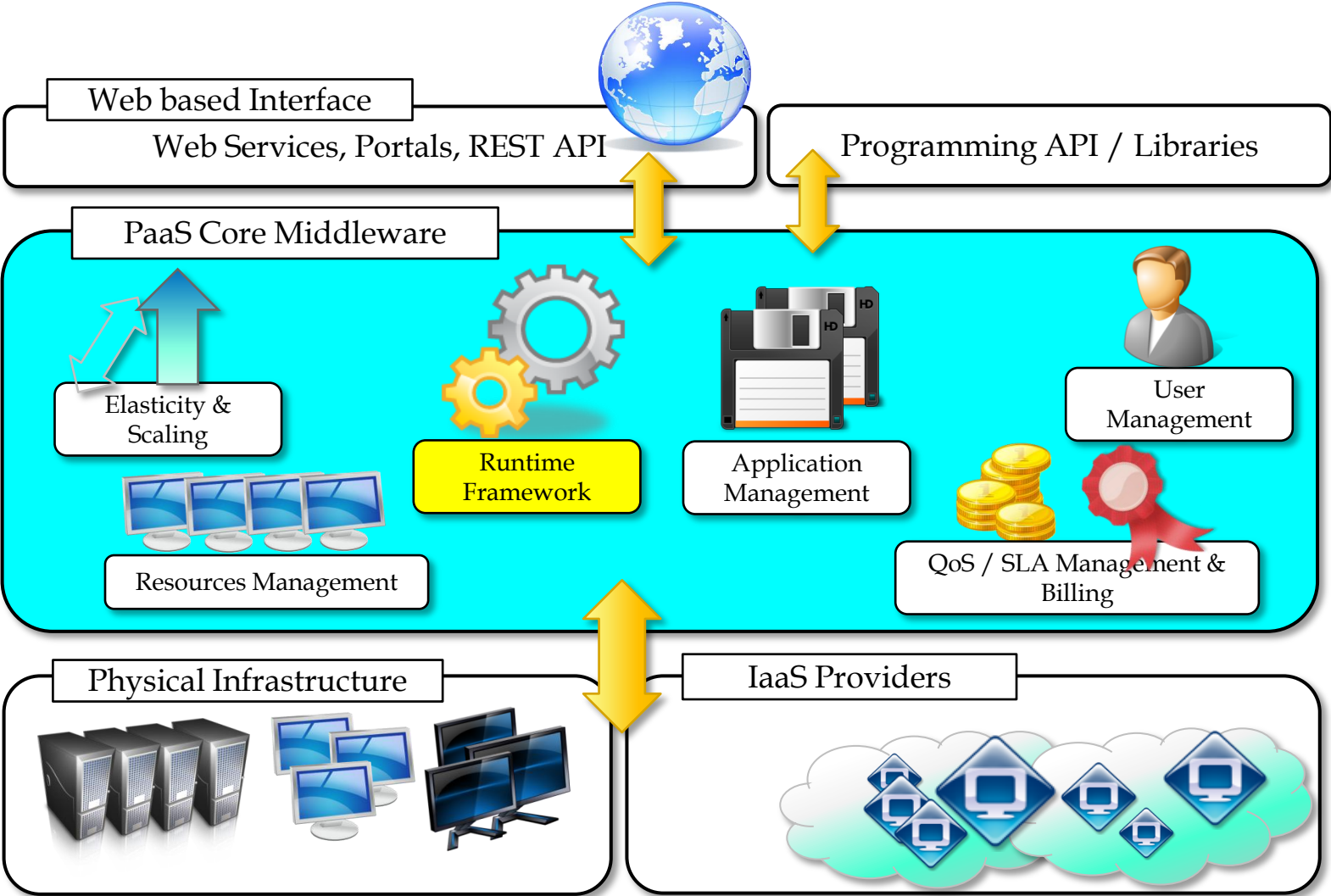
Platform as a Service (PaaS)

- **Platform-as-a-Service (PaaS)** solutions provide a development and deployment platform for running applications in the Cloud.
- They constitute the middleware on top of which applications are built.

Platform as a Service Reference Model



Platform as a Service Reference Model



PaaS Conti...

There are some essential characteristics that identify a Platform-as-a-Service solution:

- *Runtime framework*. It represents the software stack of the PaaS model.
- The runtime framework **executes end-user code** according to the policies set by the user and the provider.

PaaS Conti...

- *Abstraction* PaaS solutions are distinguished by the higher level of abstraction that they provide.
- In the case of PaaS the **focus is on the applications** the Cloud must support.

PaaS Conti...

- *Automation* PaaS environment automate the process of
 - **deploying** applications to the infrastructure,
 - **scaling** them by provisioning additional resources when needed.
- This process is performed automatically and according to the SLA made between the customers and the provider.

PaaS Conti..

- *Cloud services* PaaS offerings provide developers and architects with services and APIs helping them to simplify the creation and delivery of elastic and highly available Cloud applications.

Category of PaaS services

Category	Description	Product Type	Vendors & Products
<i>PaaS-I</i>	Runtime environment with web hosted application development platform. Rapid application prototyping.	Middleware + Infrastructure Middleware + Infrastructure	Force.com; Longjump;
<i>PaaS-II</i>	Runtime environment for scaling web applications. The runtime could be enhanced by additional components which provide scaling capabilities.	Middleware + Infrastructure Middleware + Middleware + Infrastructure Infrastructure + Infrastructure + Middleware + Infrastructure Infrastructure Middleware + Infrastructure Middleware	Google AppEngine; AppScale; Heroku; Engine Yard; Joyent Smart Platform; GigaSpaces XAP;
<i>PaaS-III</i>	Middleware and programming model for developing distributed applications in the Cloud.	Middleware + Infrastructure Middleware Middleware Middleware Middleware Middleware	Microsoft Azure; DataSynapse; Cloud IQ; Manjrasof Aneka; Apprenda SaaSGrid; GigaSpaces DataGrid;

SaaS

- The acronym SaaS was then coined in 2001 by the Software Information & Industry Association (SIIA) with the following connotation:

The acronym SaaS

- “In the **software as a service** model, the application, or service, is deployed from a centralized data center across a network – Internet, Intranet, LAN, or VPN providing access and use on a recurring fee basis. Users *rent, subscribe to, are assigned, or are granted access to* the applications from a central provider.
- Business models vary according to the level to which the software is streamlined, to lower price and increase efficiency, or value-added through customization to further improve digitized business processes.”

Cloud Services

The screenshot shows the Salesforce.com homepage. At the top, the navigation bar includes the Salesforce logo with the tagline "Success On Demand", and links for "Salesforce CRM", "The AppExchange", "Successforce.com", and "developer.force.com". A phone number "1-800-NO-SOFTWARE" is also present. Below the navigation bar is a horizontal menu with categories like "Applications", "Platform", "Services", "Customers", "Partners", "Events", "Company", "Foundation", "Customer Login", and "International Sites".

The main banner features a Starbucks coffee cup on the left and the "Salesforce ideas" logo on the right. The text reads "Let the best ideas bubble up to the top." with a "Learn More" button.

Below the banner, the headline states "The Leader in On-Demand Customer Relationship Management (CRM)" with pricing information: "As Little as \$9/User/Month for Group Edition | Full-Featured CRM Starting at \$65/User/Month". A secondary bar lists "43,600 Customers | 800+ Applications | 15 Languages".

The left sidebar highlights "CRM Software as a Service" with a "Salesforce" logo. It lists key features: "Sales Force Automation", "Marketing Automation", "Customer Service & Support", and "Partner Relationship Management".

The right sidebar contains a search bar, a "FREE TRIAL" badge for 30 days, and a "NEXT STEPS" section with buttons for "View Demo", "Contact Me", and "Resource Center". Below that is an "Announcements" section with a video link and an "Upcoming Events" section.

A large red watermark "SaaS: Salesforce.com" is overlaid on the bottom right of the page.

Cloud Services

The screenshot shows the AWS website homepage. At the top, there is a navigation menu with links for Menu, AWS Summits, Products, Solutions, Pricing, Software, Partners, Developers, Training, and Support. A search bar and a 'Create an AWS Account' button are also visible. The main banner features the text 'Gain free, hands-on experience with AWS for 12 months' and a 'Create a Free Account' button. Below the banner are four promotional tiles: 'AWS Summit' (Save the Date - Attend an AWS Summit in a city near you), 'WHAT IS CLOUD COMPUTING?' (Learn about the benefits of on-demand IT resources and pay-as-you-go pricing), 'GET STARTED WITH AWS' (Learn how to start using AWS in minutes), and 'SOFTWARE FREE TRIALS' (Try top software products from AWS Marketplace for free). The footer includes the text 'Broad & Deep Core Cloud Infrastructure Services' and a row of icons representing various AWS services.

Source: Amazon.com

IaaS: Amazon.com

IaaS & SaaS

Launch Amazon EC2 Instances

Configure the options for launching your Amazon EC2 instances



AMI: ami-ae5e6bda (j386)

Number of Hosts:

Instance Type:

Memory: 1.7 GB Virtual Cores: 1
Disk Capacity: 160 GB Architecture: 32 bits

Availability Zone:

Key Pair:

Name
✓ 123

Security Group:

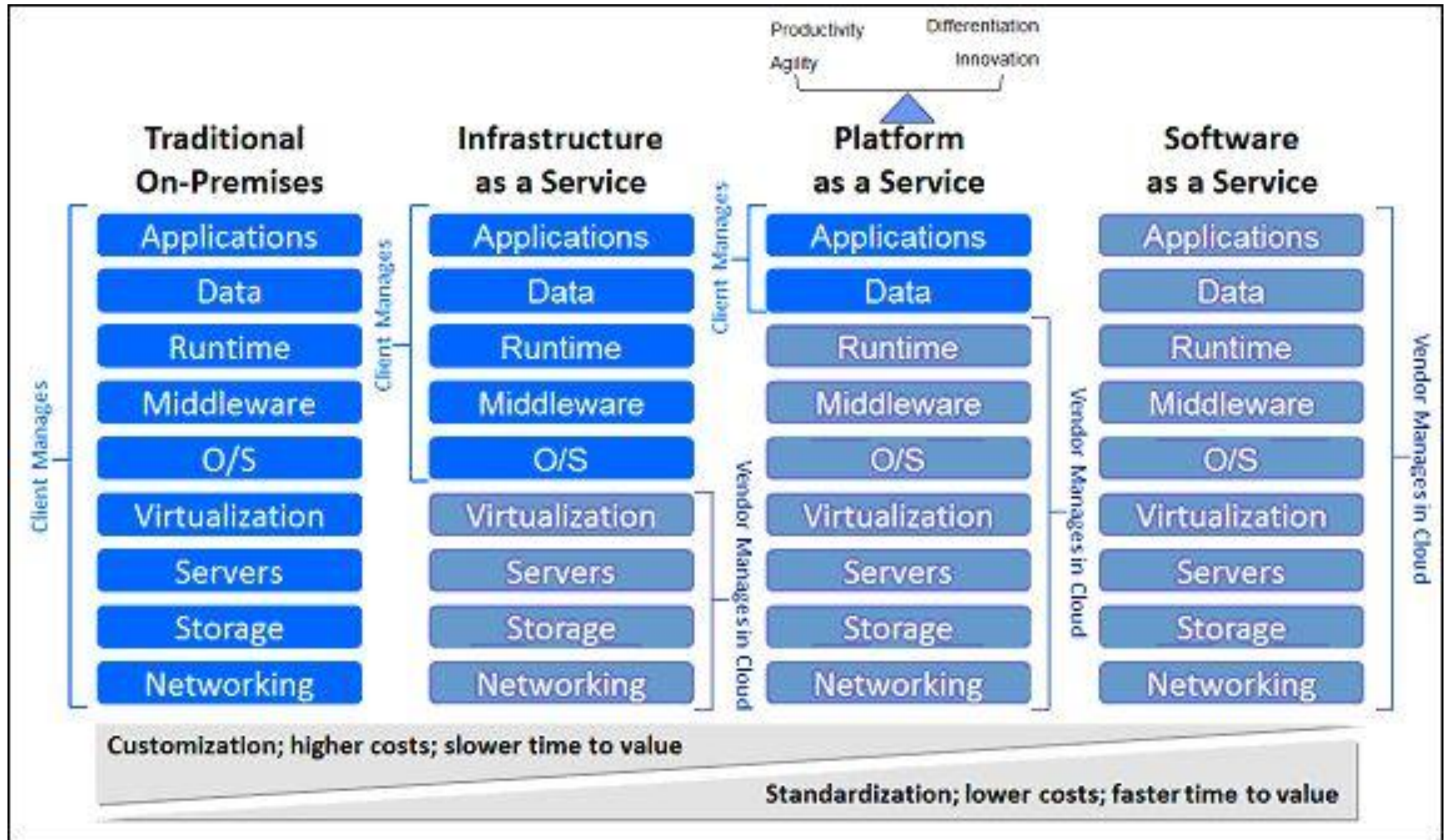
Name	Description
all	webserver
default	default group

User Data:

You will be charged the hourly rate for any instances you launch until you successfully shut them down.

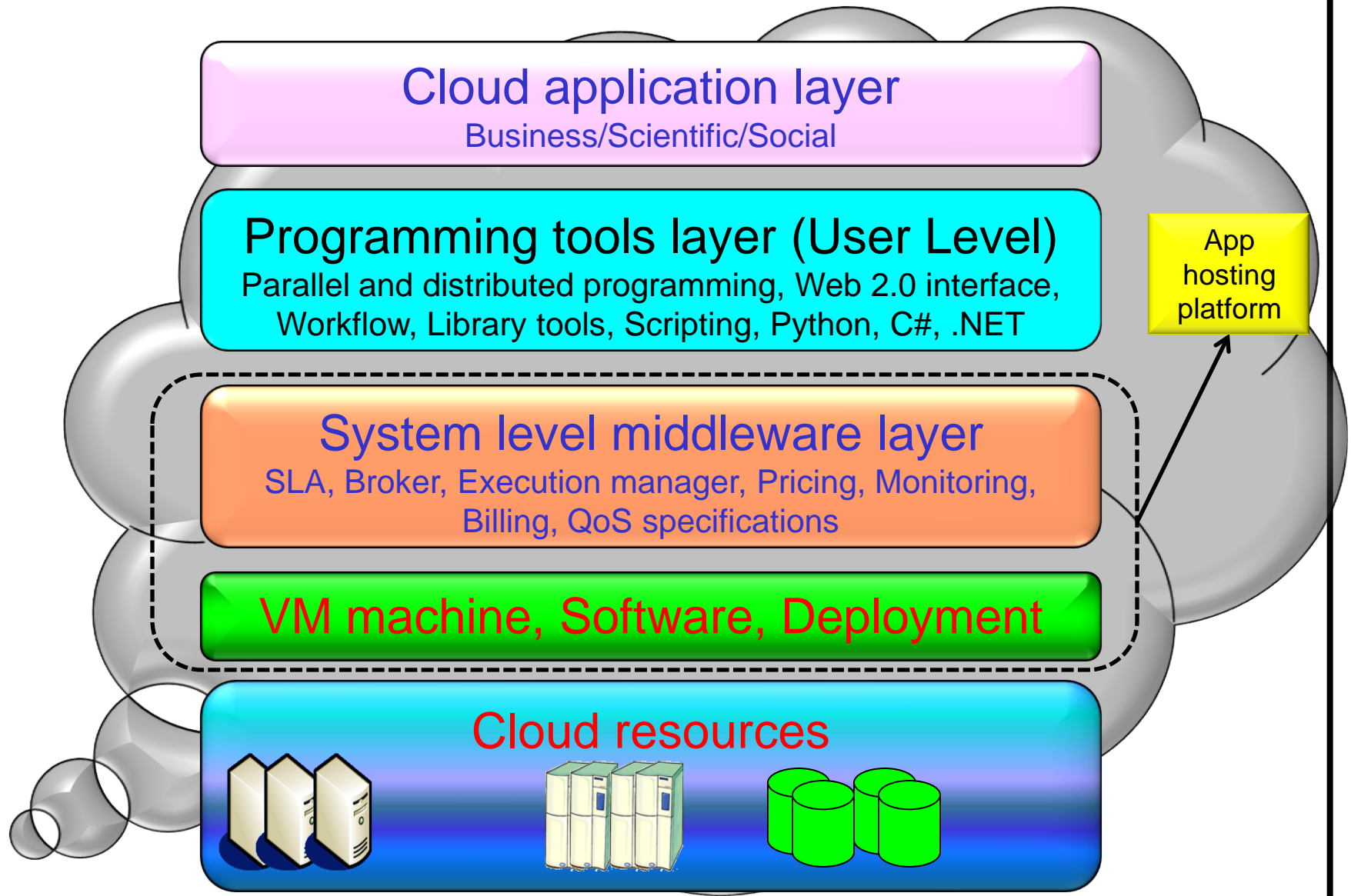


Source: Internet



Source: Internet

Layered architecture of Clouds



Entities of the clouds

Users/Brokers: Users or brokers (acting on their behalf) submit service requests from anywhere in the world to the Data Center and Cloud to be processed.

Entities of the clouds

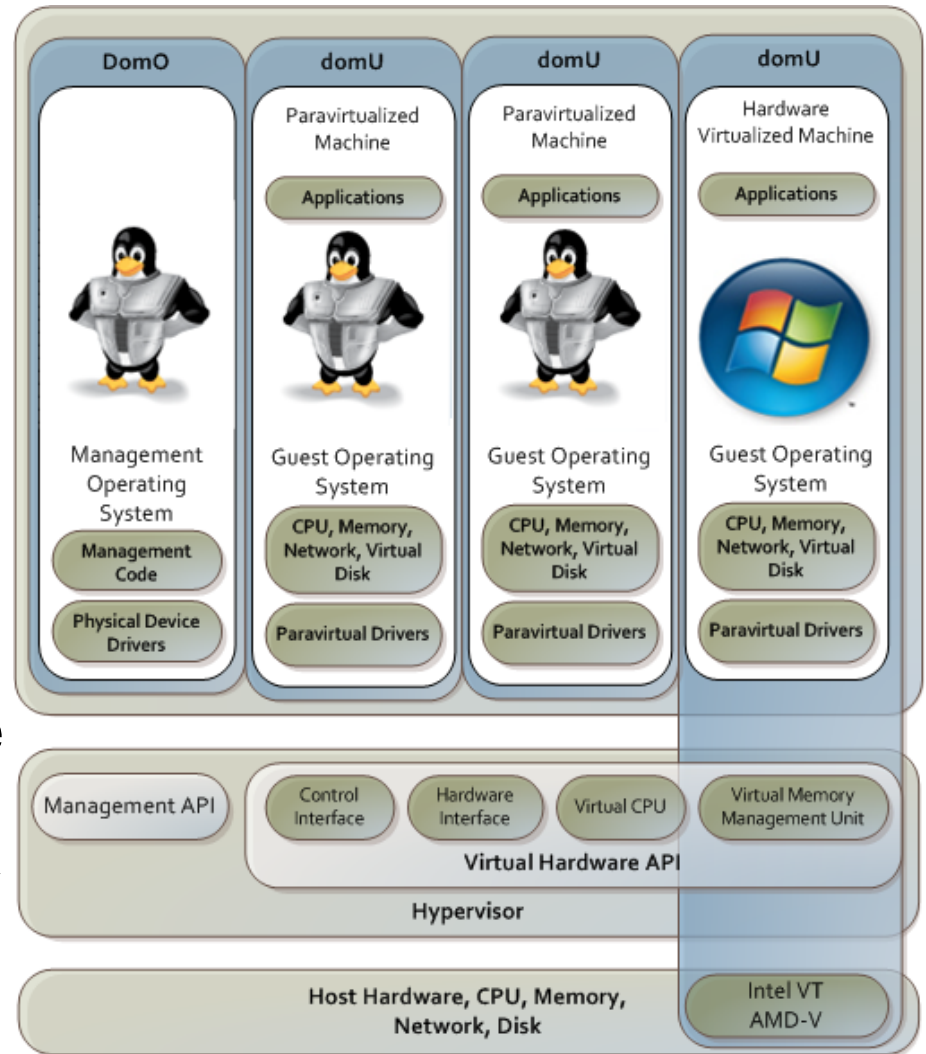
SLA Resource Allocator: *The SLA Resource Allocator acts as the interface between the Data Center/Cloud service provider and external users/brokers. It requires the interaction of the following mechanisms to support SLA-oriented resource management:*

- *Service Request Examiner and Admission Control*
- *Pricing*
- *Accounting*
- *VM Monitor*
- *Dispatcher*
- *Service Request Monitor*

Entities of the clouds

VMs:

- Multiple VMs on a single physical machine to meet service requests
- Multiple VMs can concurrently run applications based on different operating system environments on a single physical machine
- every VM is completely isolated from one another on the same physical machine.



Clouds based on Ownership and Exposure

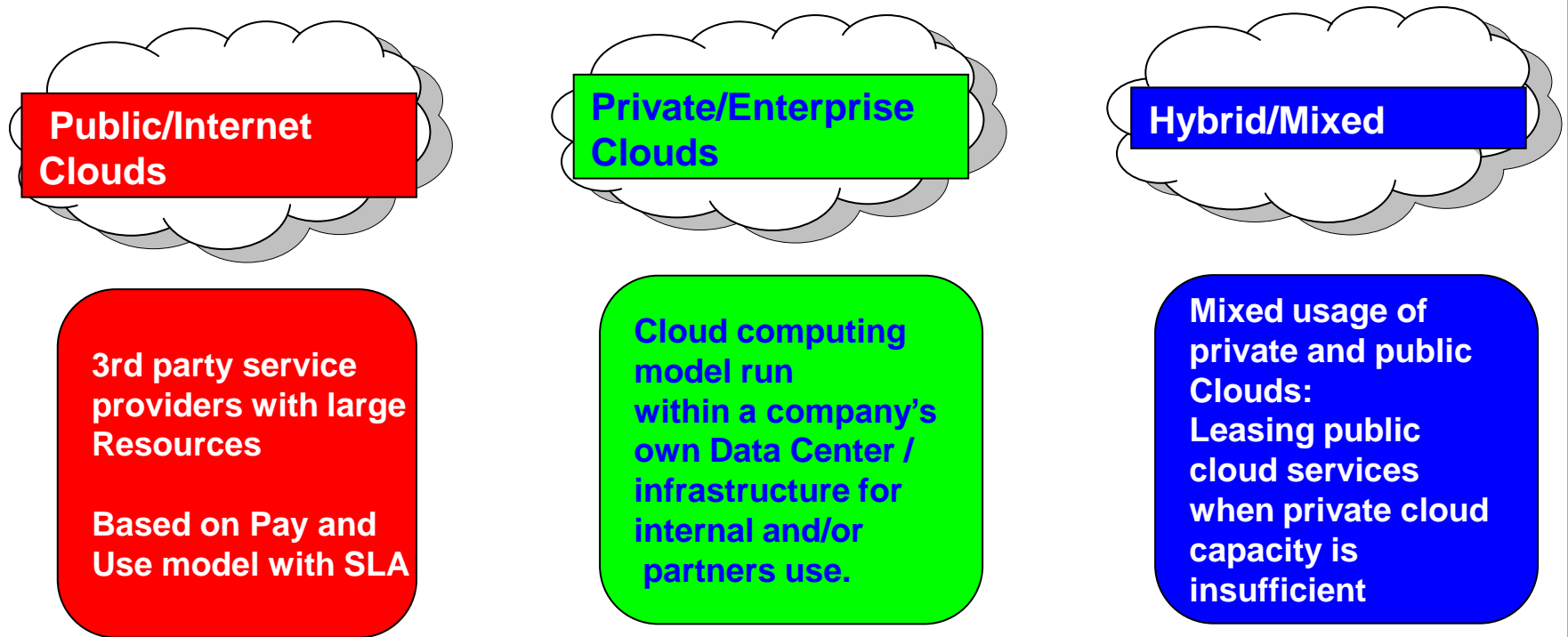


Figure Courtesy: Raj Buyya

Types of Clouds

- A more useful classification is given according to the administrative domain of a Cloud:
 - *Public Clouds*: the Cloud is open to the wide public;
 - *Private Clouds*: the Cloud is implemented within the private premises of an institution and generally made accessible to the members of the institution or a subset of them;

Types of Clouds

- A more useful classification is given according to the administrative domain of a Cloud:
 - *Hybrid or Heterogeneous Clouds*: the Cloud is a combination of the two previous solutions and most likely identifies a Private Cloud that has been augmented with resources or services hosted in a Public Cloud;

Types of Clouds

- A more useful classification is given according to the administrative domain of a Cloud:
 - *Community Clouds*: the Cloud is characterized by a multi-administrative domain, involving different deployment models (public, private, and hybrid), and it is specifically designed to address the needs of a specific industry.

Public Cloud

- Realization of the canonical view of Cloud computing
- Any customer can easily sign-in with the Cloud provider
- Ability to grow or shrink according to the need of the related business
- *A fundamental characteristic of Public Clouds is multi-tenancy*
- A Public Cloud can offer any kind of service: infrastructure, platform, or applications. For example, Amazon EC2 is a Public Cloud

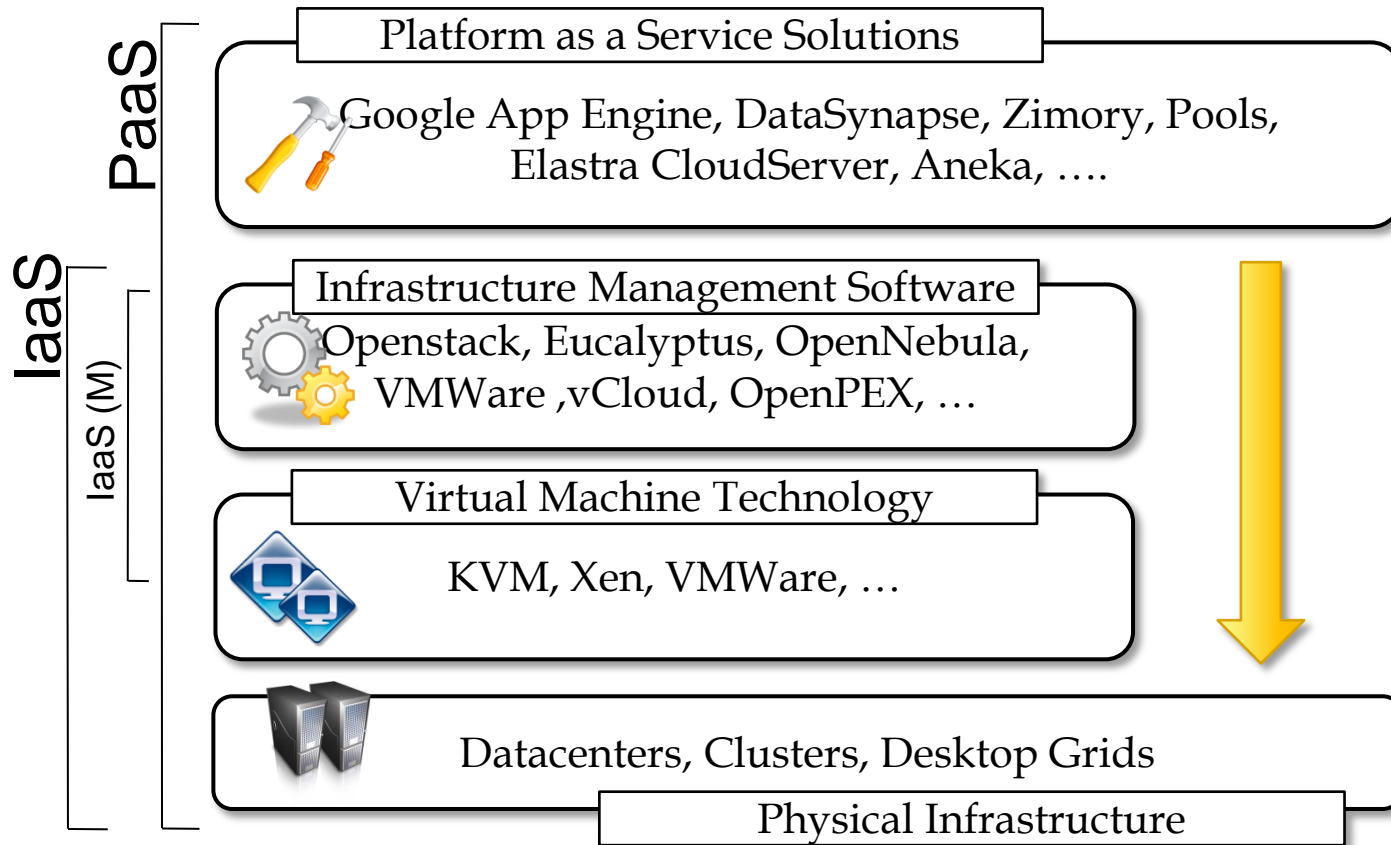
Private Clouds

- In Public Cloud the provider is in control of the infrastructure and customers' core logic and sensitive data results in *loss of control*
- This could open the way to other problematic situations
- Private Clouds are virtual distributed systems that rely on a private infrastructure and provide internal users with dynamic provisioning of computing resources

Private Clouds

- Possibility of testing applications and systems at a comparatively lower price rather than Public Clouds before deploying them on the public virtual infrastructure
- *Customer information protection*
- *Infrastructure ensuring Service Level Agreements (SLAs)*
- *Compliance with standard procedures and operations*

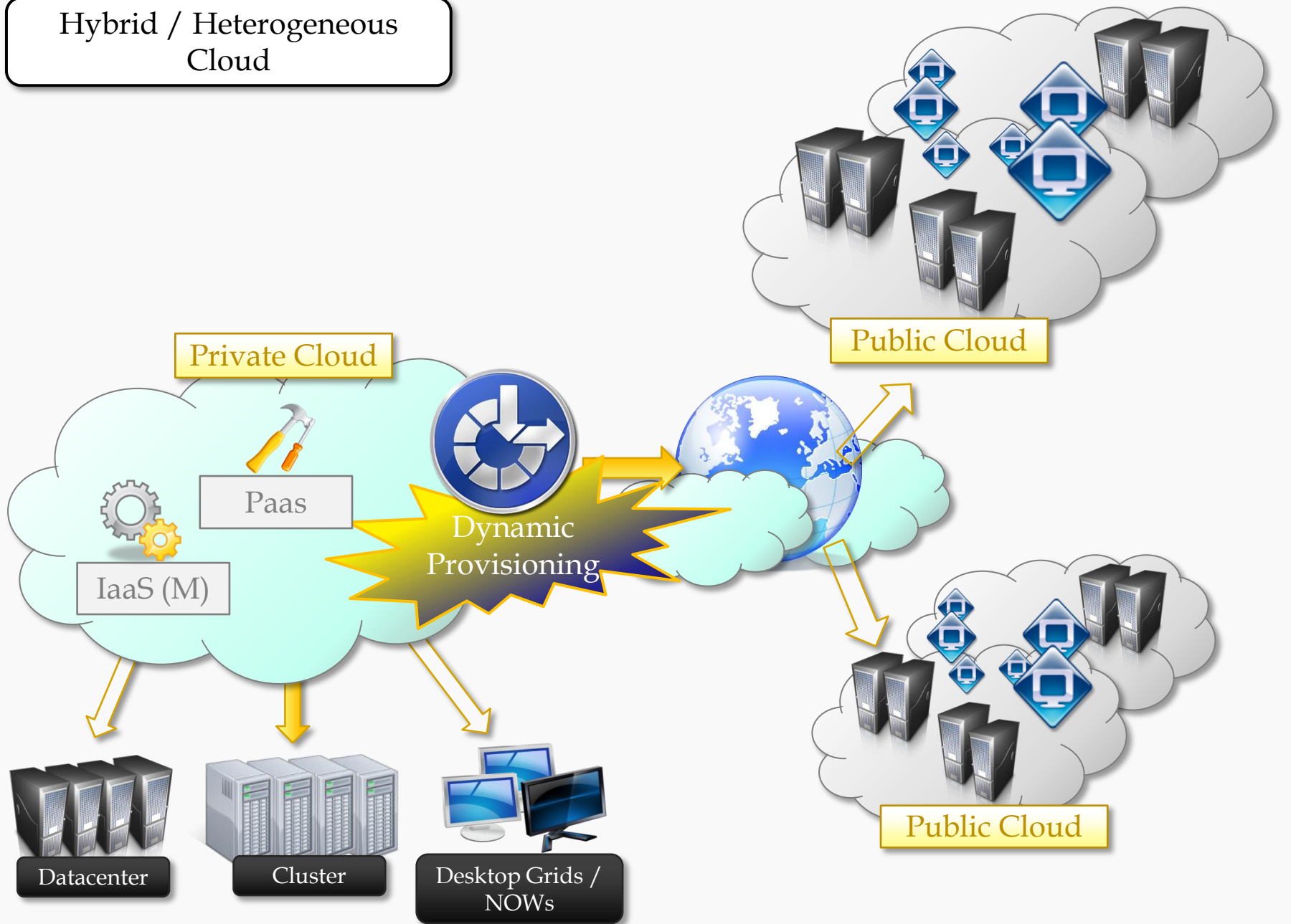
Private Clouds Hardware and Software Stack



Hybrid Clouds

- A hybrid cloud is an environment in which an organization provides and manages some resources **in-house** and has others provided **externally**. Example Amazon S3 for archived data but continue to maintain in-house storage for operational customer data.
- Take advantage of the **scalability** and **cost-effectiveness** that a public cloud computing environment offers without exposing mission-critical applications and data to third-party vulnerabilities

Hybrid / Heterogeneous Cloud

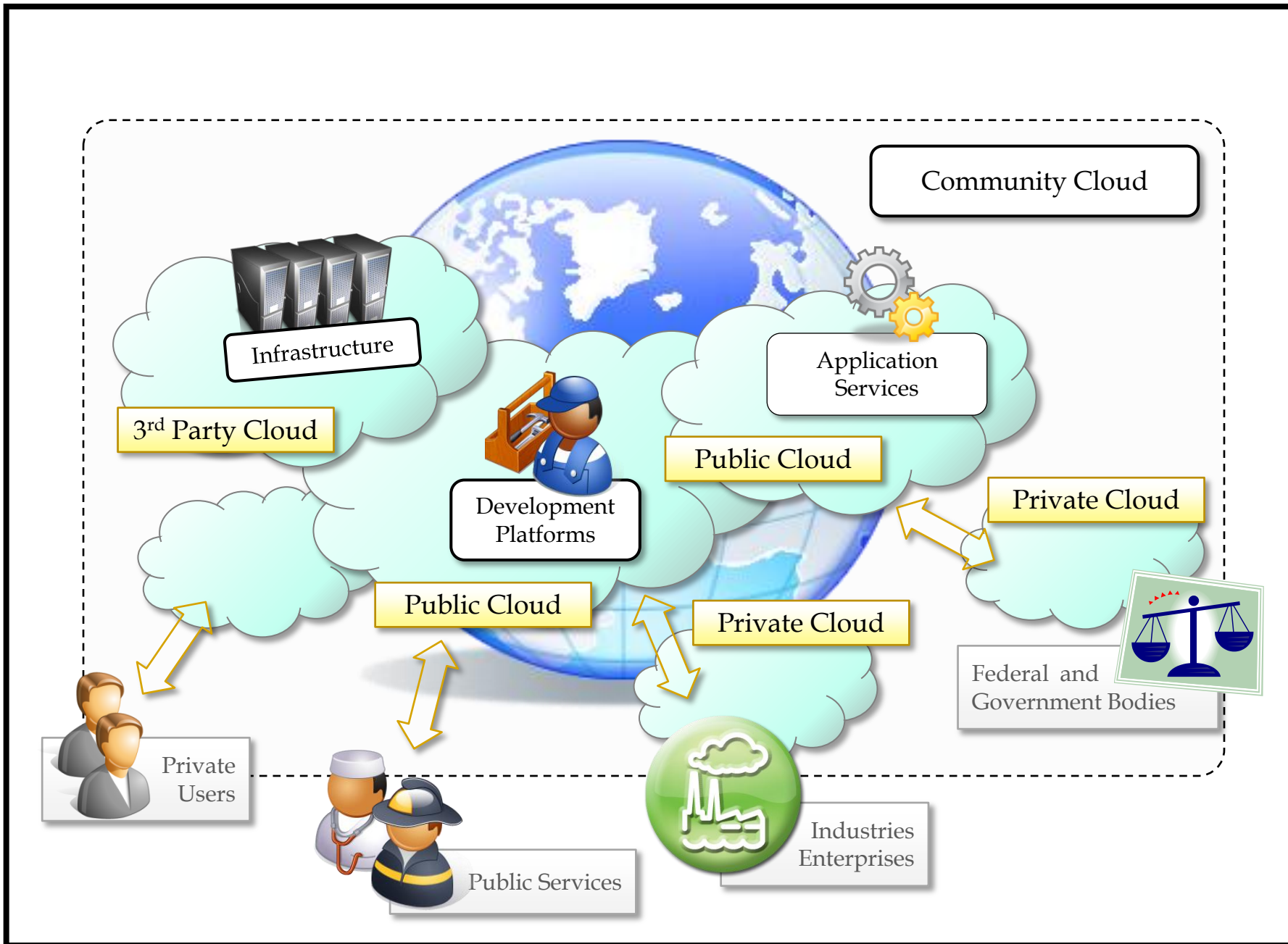


Community Clouds

- Community Clouds are distributed system constituted by integrating the services of different Clouds to address the specific needs of an industry, a community, or a business sector

Community Cloud

- *The infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g. mission, security requirements, policy, and compliance considerations).*
- *It may be managed by the organizations or a third party and may exist on premise or off premise.*



- The benefits of these Community Clouds are the following
 - *Openness*
 - *Community*
 - *Graceful failures*
 - *Convenience and control*
 - *Environmental sustainability*

Example Cloud Platforms

Amazon Elastic Compute Cloud (EC2) (IaaS)

- It provides a virtual computing environment that enables a user to run Linux/Windows etc. applications.
- The user can **either create (using AWS)** a new Amazon Machine Image (AMI) containing the applications, libraries, data and associated configuration settings, **or select** from a library of globally available AMIs.
- The user then **needs to upload** the created or selected AMIs to Amazon Simple Storage Service (S3), before they can start, stop, and monitor instances of the uploaded AMIs.

Ref.: Amazon Elastic Compute Cloud (EC2). <http://www.amazon.com/ec2/>

Ref.: Amazon Simple Storage Service (S3). <http://www.amazon.com/s3/>

Example Cloud Platforms

Google App Engine (PaaS)

- Google App Engine allows a user to run web applications written using the **Python, Java, PHP, Go, Node.js, .NET, Ruby** programming languages.
- It also supports (APIs) for the data store:
 - **App Engine Datastore**
 - **Google cloud SQL**
 - **Google cloud store**
- Google Accounts, URL fetch, image manipulation, and email services. Security is provided with ***The Sandbox***.

Ref.: Google App Engine. <http://appengine.google.com>

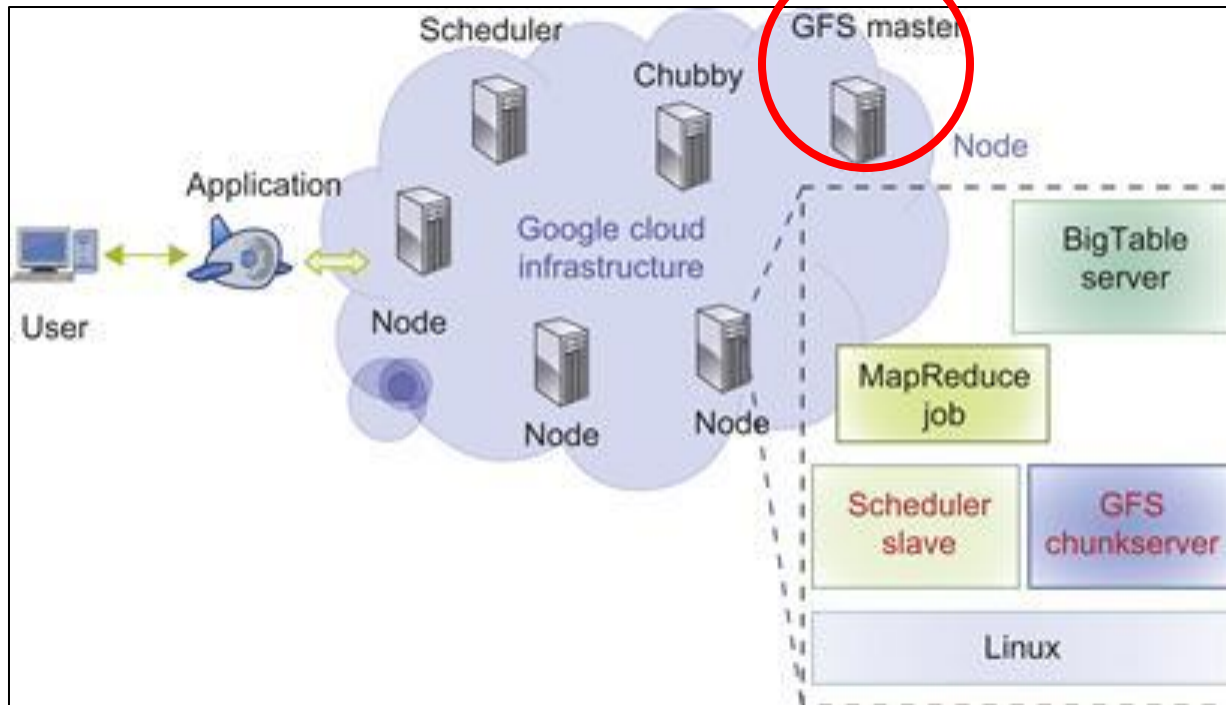
Example Cloud Platforms

Google App Engine (PaaS)

- It also provides a web-based Administration Console for the user to easily manage its running web applications.
- Currently, Google App Engine is free to use with up to 1TB of storage and about 5 million page views per month.
- *Compute Engine (IaaS) and Container Engine (Virtual Cluster)*

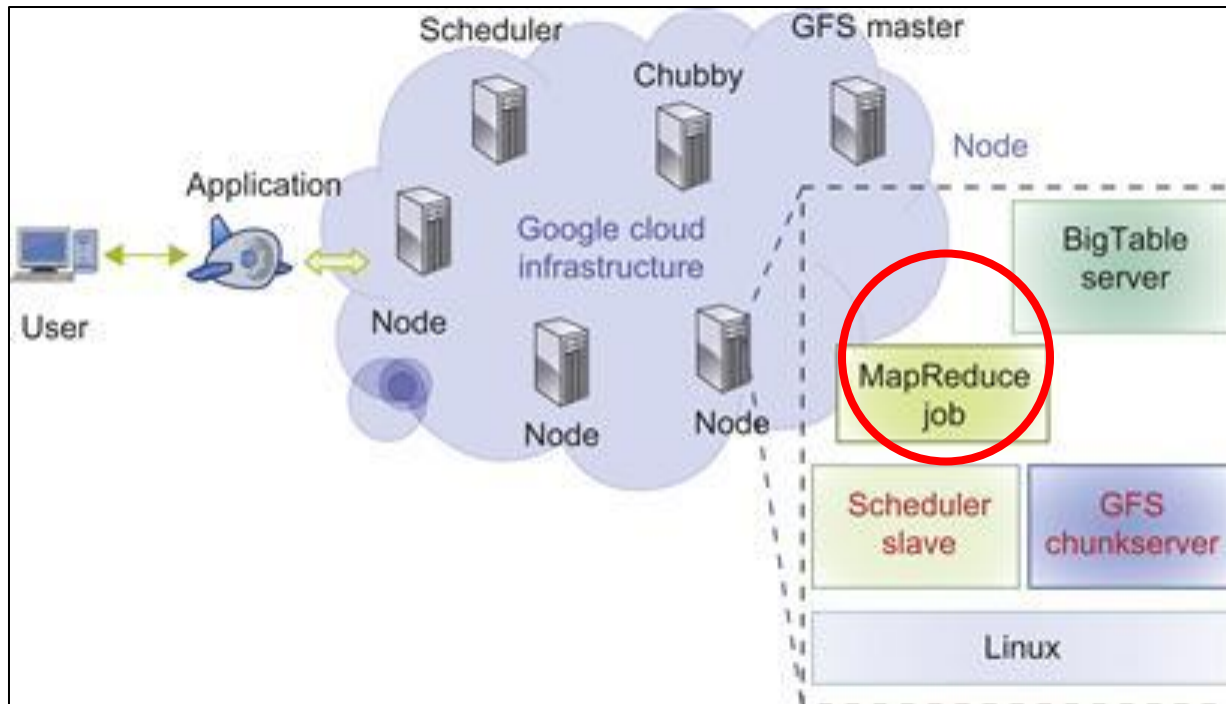
Google App Engine. <http://appengine.google.com>

Google cloud platform and major building blocks



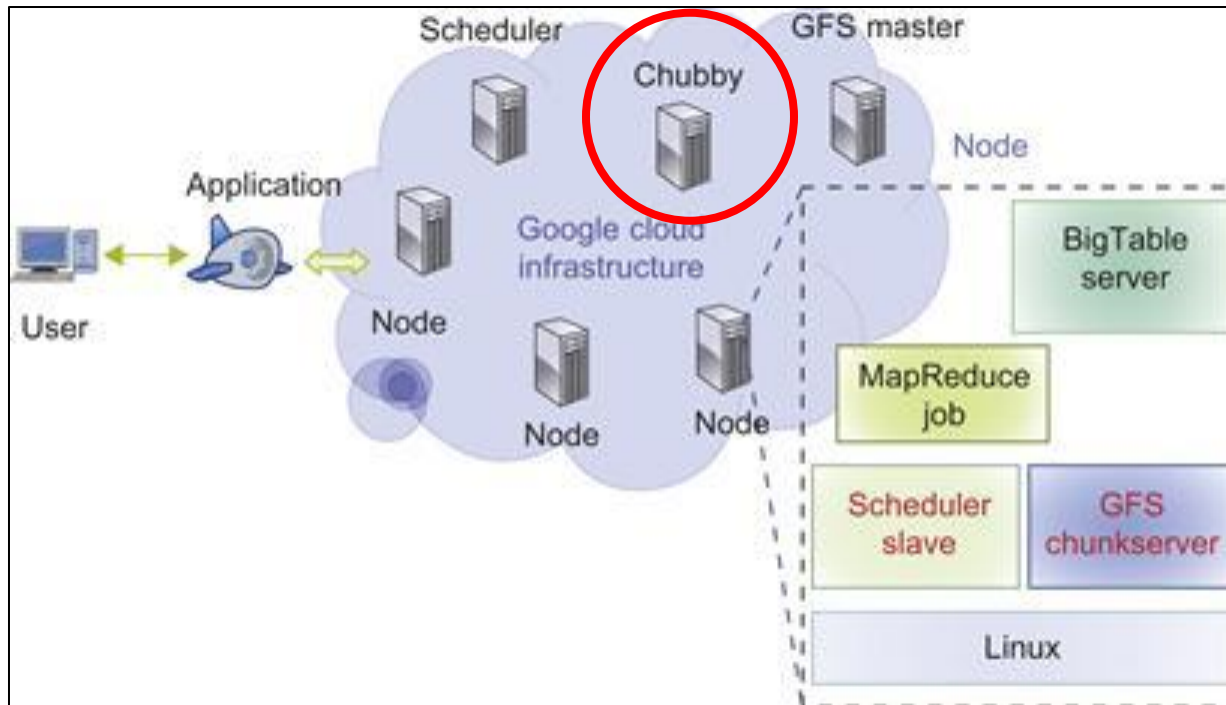
- ***GFS*** is used for storing large amounts of data.

Google cloud platform and major building blocks



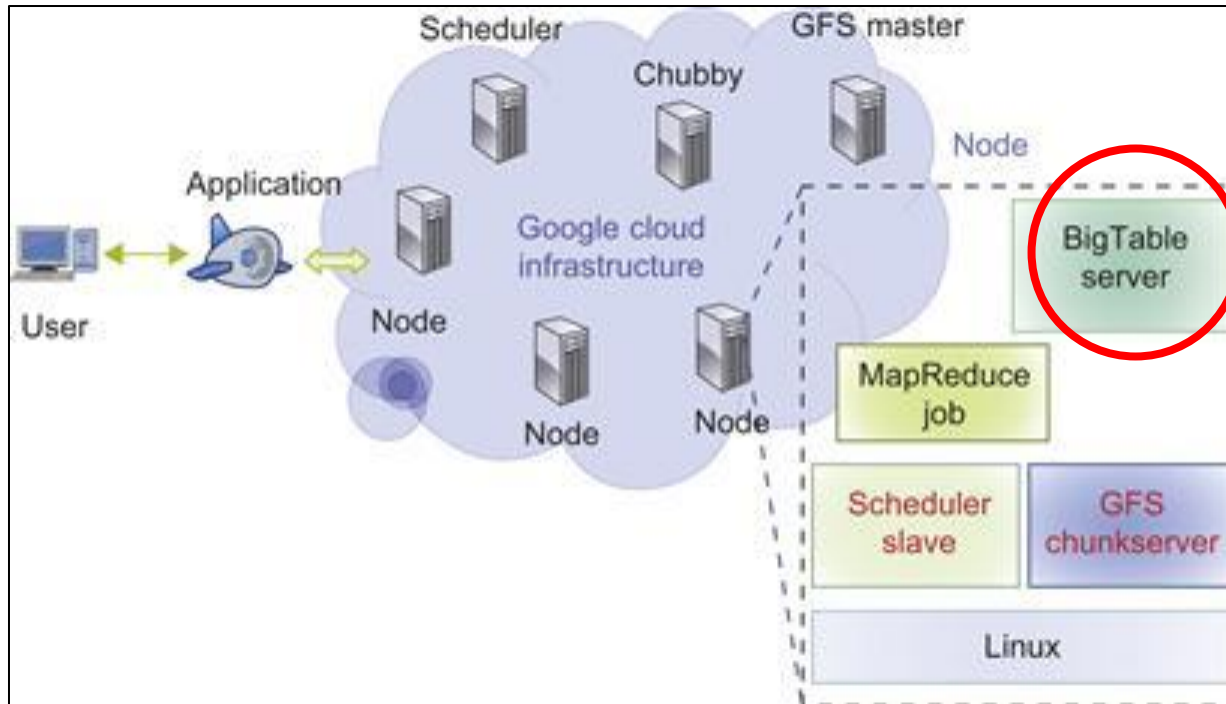
- *MapReduce* is for use in application program development.

Google cloud platform and major building blocks



- *Chubby* lock service for loosely-coupled distributed systems. It is designed for coarse-grained locking.

Google cloud platform and major building blocks



- *BigTable* offers a storage service for accessing un-structured data.

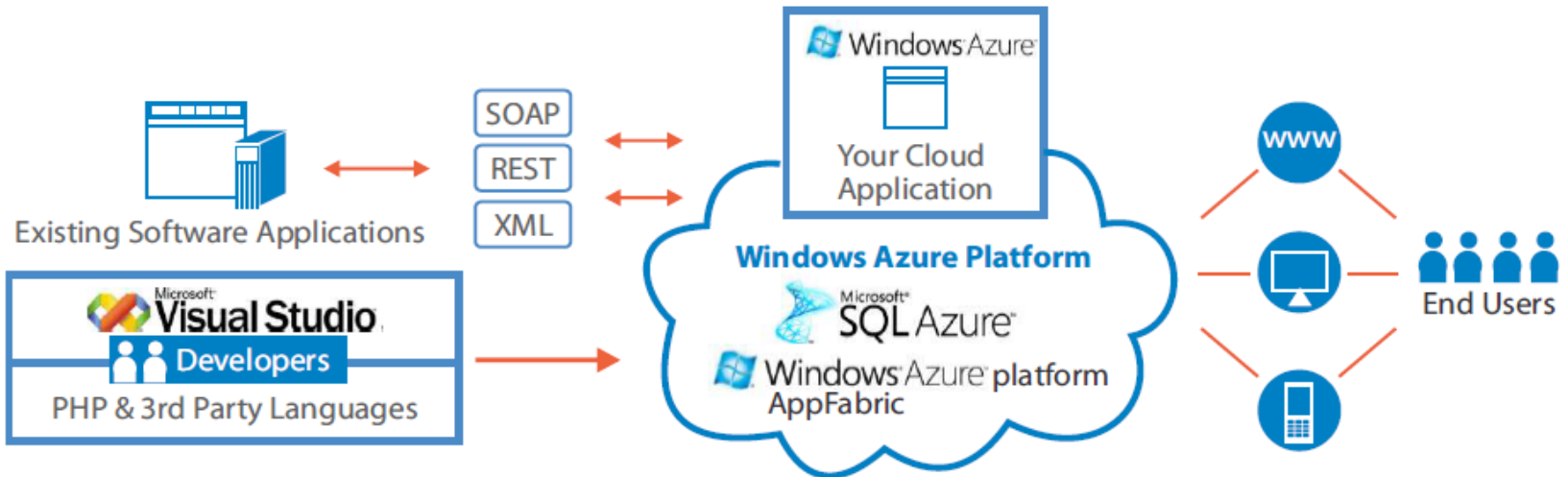
Example Cloud Platforms

Microsoft Azure (PaaS)

- Aims to provide an integrated development, hosting, and control Cloud computing environment so that developers can create, host, manage, and scale both Web and non-web applications through Microsoft data centres
- It supports a comprehensive collection of proprietary development
- It supports Web APIs such as SOAP and REST to allow software developers to interface between Microsoft or non-Microsoft tools and technologies.

Microsoft Azure. <http://www.microsoft.com/azure/>

Windows Azure Platform Footprint



Windows Azure

- **Web Role** is customized for web application programming and supported by Internet Information Services (IIS 7, web server by MS).
- **Worker Role** is used for performing the background process for the web role. Message queue end points are configured to generate programs in an asynchronous mode.

Windows Azure

- **VM Role runs an image** (a VHD) of a *Windows Server 2008 R2* virtual machine.
- First VHD is created using on-premises Windows Server machine, then uploaded to Windows Azure.
- Customers can configure and maintain the OS and use Windows Services, scheduled tasks etc. in the VM role.

Windows Azure

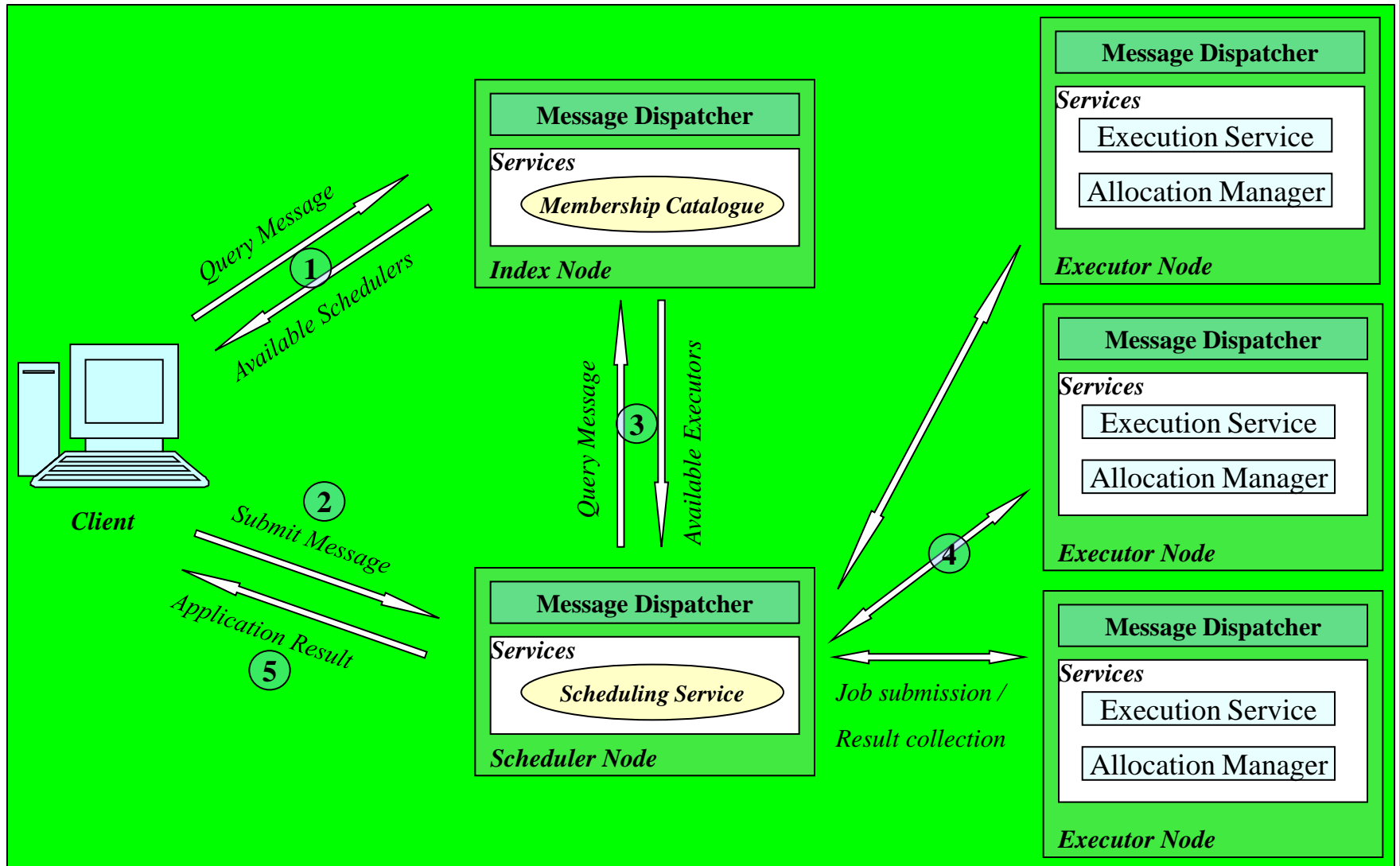
- Once it is stored in the cloud, the VHD can be loaded on-demand into a VM role and executed. The VHD can be used as the base image for all instances of a VM Role.
 - OS disk
 - Temporary disk (data loss in case of VM failure, size 20 ~ 605 GB)
 - Data disk

Example Cloud Platforms



Aneka (PaaS)

- A 3rd generation and SOA-based Enterprise PaaS Cloud Developed by GRIDS Lab at the University of Melbourne (Australia) In partnership with Manjrasoft
- Supporting multiple models for programming parallel and distributed applications (thread, task or MapReduce).
- Flexible and extensible architecture.
- Enterprise-class QoS Applications can negotiate for the required capability


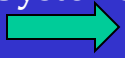
Program Execution in Aneka



Comparison of Cluster, Grid and Cloud

Properties  	Cluster	Grid	Cloud
Elements	Commodity computers	Servers, clusters	Grids, Network Storage
Scalability	Realistic if < 500	1000's	User's requirement generally 100 to 1000
OS	Any one of these (Linux, Solaris, Windows etc.)	Any standard OS highly dominated by Linux	Hypervisor on multiple OS runs
Network	LAN, dedicated (tightly coupled) high bandwidth	Mostly Internet	Dedicated, priority is high bandwidth
Security	Limited, username/password	Using public key cryptology, mapping user to an account	High, each user has virtual machine
Resource discovery	Membership service	Centralized Indexing, decentralized info service	Membership service
Service Negotiation	Limited	SLA based	SLA based (more price oriented)

Comparison of Cluster, Grid and Cloud

Properties   Systems	Cluster	Grid	Cloud
Resource management	Centralized	Mostly Decentralized	Both Centralized and distributed
Resource scheduling	Centralized	Distributed, Global and local schedulers are different	Both Centralized and distributed
Standard Interface	Virtual Interface Architecture	OGSA based, now majority SOA	SOAP and REST
Single System Image	Yes	No	Yes
Pricing of Service	Partially but limited	Objective was collaborative use of resources	Highly business model
Important Systems/Examples	Beowulf Cluster	Globus/Gridbus	EC2, GoogleApp
Research opportunities	Yes	Yes	Limited