

# Web Technologies

## Lecture 13

### Introduction to cloud computing

# Cloud Computing

- **NIST:**

“a model for enabling **ubiquitous**, 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.”

- **Gartner:**

“a style of computing in which **scalable** and **elastic IT-enabled** capabilities are **delivered as a service** using **Internet** technologies.”

- **Webster:**

“the practice of **storing** regularly used computer **data** on multiple servers that can be **accessed through the Internet**.”

# On demand pay per use

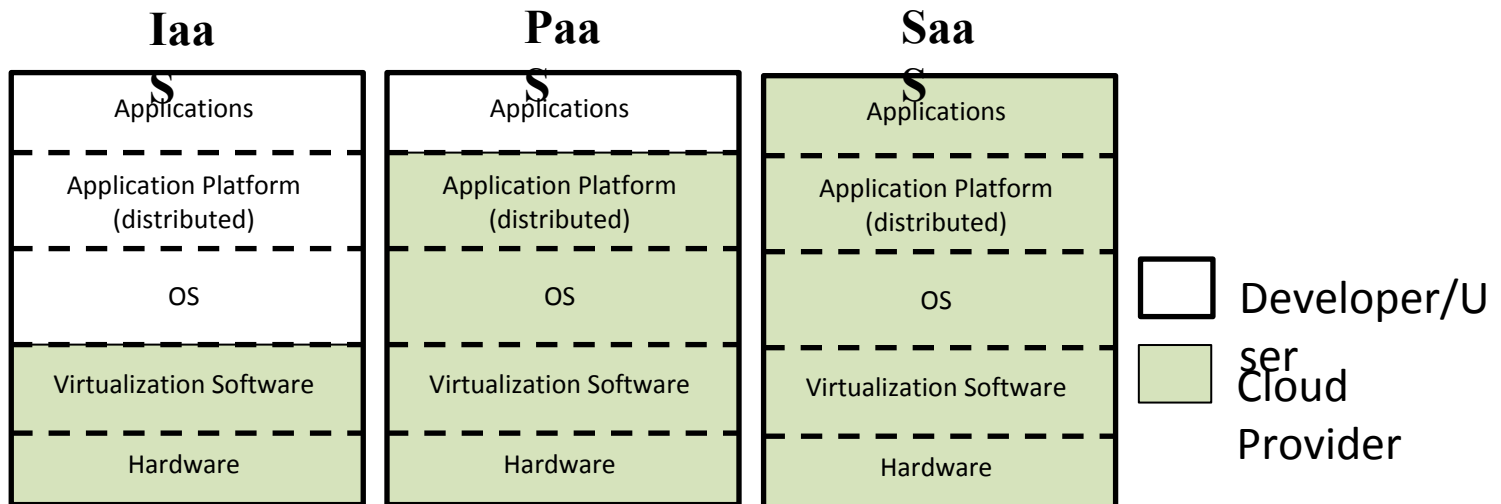
- On demand **self-service**
  - A consumer can provision computing services such as server resources, network, storage automatically as required
- Geographically **distributed** and broad network access
  - The services are geographically distributed and are accessible broadly through internet and hence can be used with thick or thin clients (such as mobile phones)
- Multi-tenant **resource pooling**
  - The resources are shared among multiple clients using the cloud for better utilization of the underlying infrastructure. The allocation of resources is transparent to the consumer
- Pay per use
  - Hourly (Amazon, Azure)
  - Per minute (Google)

# Elasticity

- Rapid **elasticity**
  - The consumer can rapidly request or release resources based on their requirements. This is useful in quickly scaling out to changing demands
- Measured service (**utility computing**)
  - Different metering mechanisms are in place to monitor and track individual usages for services such as compute, network, storage and any other specialized services. The consumer only pays for the actual services used (pay-per-use model)

# Service models

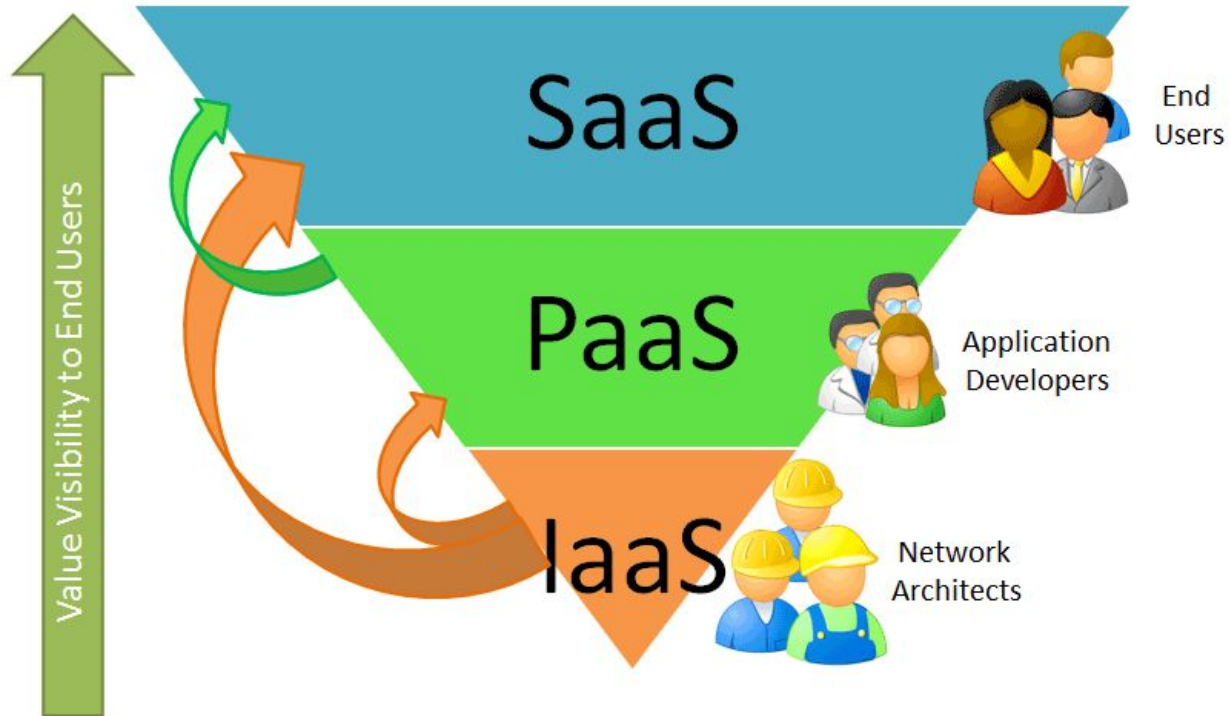
- Infrastructure as a Service (IaaS) - utility computing
  - datacenter as a service
  - client provisions processing, storage, networks where she/he can run arbitrary software s.t. OS or applications
  - full control of the infrastructure through virtualized resources
  - **Example:** Amazon EC2



# Service models (2)

- Platform as a Service (PaaS)
  - developer point of view
  - deploy applications on the cloud using programming languages, libraries, and tools provided by the cloud provider
  - no management control of the infrastructure
  - **Example:** Google App Engine, Microsoft Azure
- Software as a Service (SaaS)
  - end user point of view
  - use existing applications deployed on the cloud
  - software experiences are delivered through the Internet
  - **Example:** Google Drive, Flickr, Gmail

# Cloud stack



# Deployment models

- Private clouds
  - Single organization owned, managed, and operated cloud infrastructure
  - Community clouds
  - Infrastructure provisioned for exclusive use by a specific community of consumers from organizations that have
  - Shared concerns
  - Owned, managed, and operated by one or more organizations in the community



# Deployment models (2)

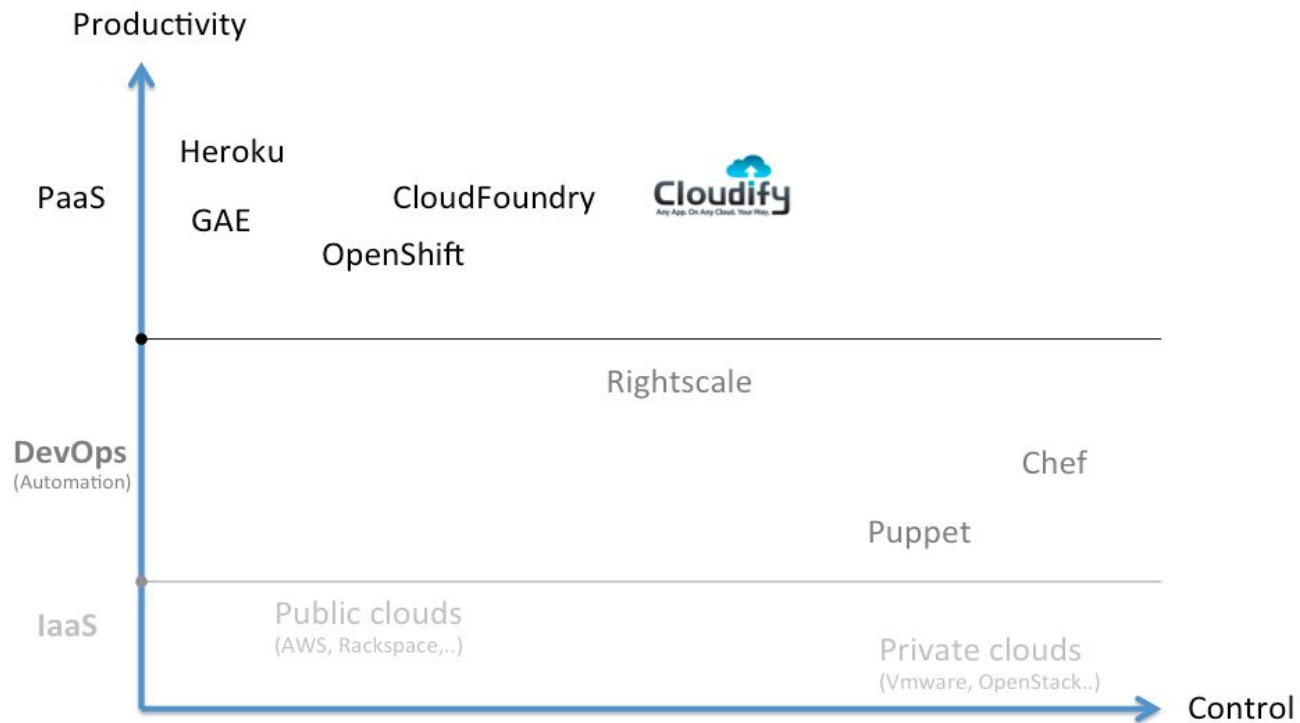
- Public clouds
  - cloud infrastructure provisioned for public use
  - owned, managed, and operated by a business, academic or government institution, or a combination of them
  - usually accessible following a pay-per-use billing model
- Hybrid clouds
  - composition of two or more private/community/public clouds that remain unique entities but are bound by standards or proprietary technology

# Control vs. productivity

- Different service models (IaaS, PaaS) provide different levels of control and productivity in terms of management overhead and administration requirements
- IaaS provide greater control since everything from the OS to platform to application is under developer's control
- PaaS give higher productivity since the details of the underlying platform are completely hidden and transparent to the user  
(e.g. handling scalability or VM lifecycle management)
- Public vs. Private vs. Hybrid cloud deployments also imply different level of control over the infrastructure as well as data and computation

# Control vs. productivity

## THE CLOUD STACK



# Cloud concepts: virtualization

- Cloud computing main aspect: elastic on demand
  - pay as you go
  - use as much as you want whenever you want
- These notions are practical only if we have
  - lot of flexibility
  - efficiency in the back-end
- These are readily available in Virtualized Environments and Machines

# Virtualization

- The creation of a virtual (rather than actual) version of something, such as an operating system, a server, a storage device or network resources
- Allows sharing of physical resources among multiple users (tenants)
- Allows deployment of hardware agnostic software
- Allows easy configuration of virtual machine images and quick deployment of large number of services

# Virtualized vs. traditional computing

Traditional computing stack

Virtualized computing stack

# Hypervisor/VMM

Software layer which:

- Allows multiple guest OSs (Virtual Machines) to run simultaneously on a single physical host
- Provides a hardware abstraction to the running guest OS and efficiently multiplexes underlying hardware resources

# Multiprogramming vs. virtualization

- **Multi Programming** – Each Process thinks it has complete control on all of the resources
  - Virtual Memory
  - CPU Sharing
- **Virtualization** – OS assumes control of the entire underlying infrastructure through a hypervisors/VMM



# Multiprogramming vs. virtualization (2)

Multi Programming	Virtualization
1. CPU shared among processes	1. CPU shared among OSs
2. Memory shared using Page Tables	2. Memory shared using more indirections: Multiple Page Tables
3. Process knows it is being managed (system calls)	3. OS may/ may not know it is being used

# Amazon Web Services (AWS)

- First Public Cloud (launched in 2006)
- Collection of on-demand pay-per-use computing services
- Solutions in various service models:
  - IaaS: EC2, S3, ELB, Autoscaler
  - PaaS: Elastic Beanstalk, EMR
  - SaaS: Cloud Search, Elastic Transcoder
- Other services:
  - Networking: DNS, CDN
  - Databases: relational, noSQL, memcache
  - Scripted deployment