

Cloud Computing

JEONBUK NATIONAL UNIVERSITY

Contents

A. What is Cloud Computing?

B. Study Roadmap

1 What is Cloud Computing?

JEONBUK NATIONAL UNIVERSITY

- **Cloud computing** is the delivery of hosting services that are provided to a client over the Internet.
- Enable large-scale services without up-front investment.

•NIST Definition of Cloud Computing

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.

- **Cloud computing is Utility Computing**

Cloud services are controlled and monitored by the cloud provider through a pay-per-use business model.

•An ideal cloud computing platform

- efficient in its use of resources
- elastic
- self-managing
- highly available and accessible

- **Resource efficiency**

computing and network resources are pooled to provide services to multiple users. Resource allocation is dynamically adapted according to user demand.

- **Elasticity**

computing resources can be rapidly and elastically provisioned to scale up, and released to scale down based on consumer's demand.

- **Self-managing services**

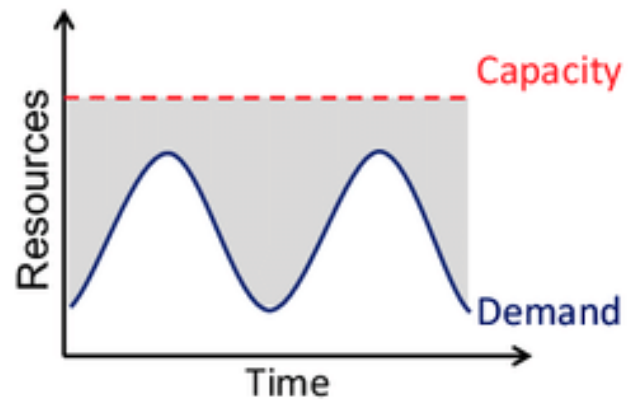
a consumer can provision cloud services, such as web applications, server time, processing, storage and network as needed and automatically without requiring human interaction with each service's provider

- **Accessible and highly available**

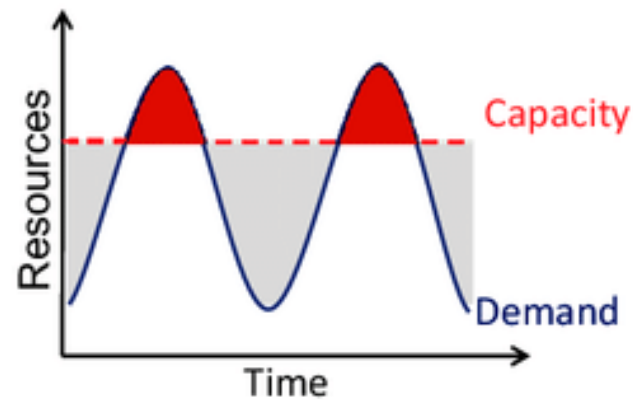
cloud resources are available over the network anytime and anywhere and are accessed through standard mechanisms that promote use by different types of platform (e.g., mobile phones, laptops, and PDAs).

- over-provisioning, under-provisioning

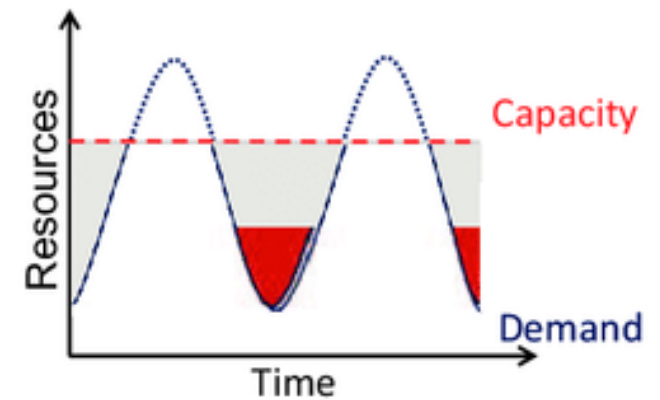
Over-provisioning




Under-provisioning

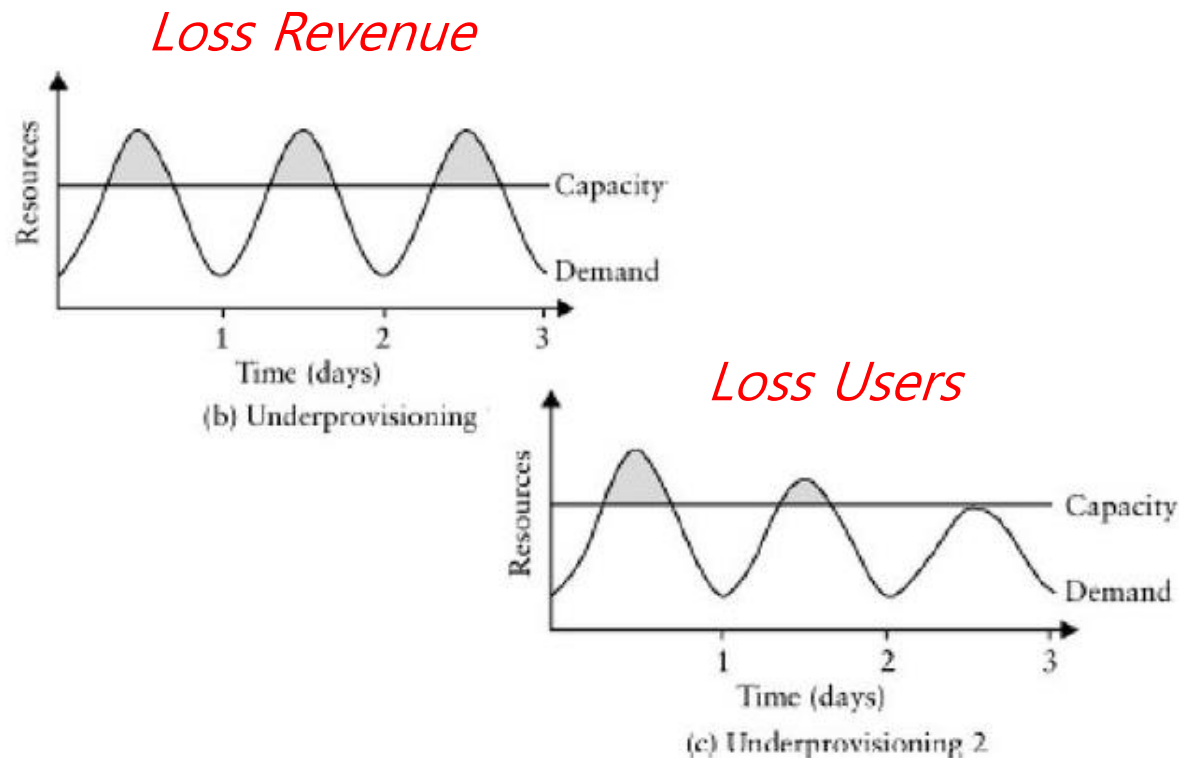


Delayed allocation

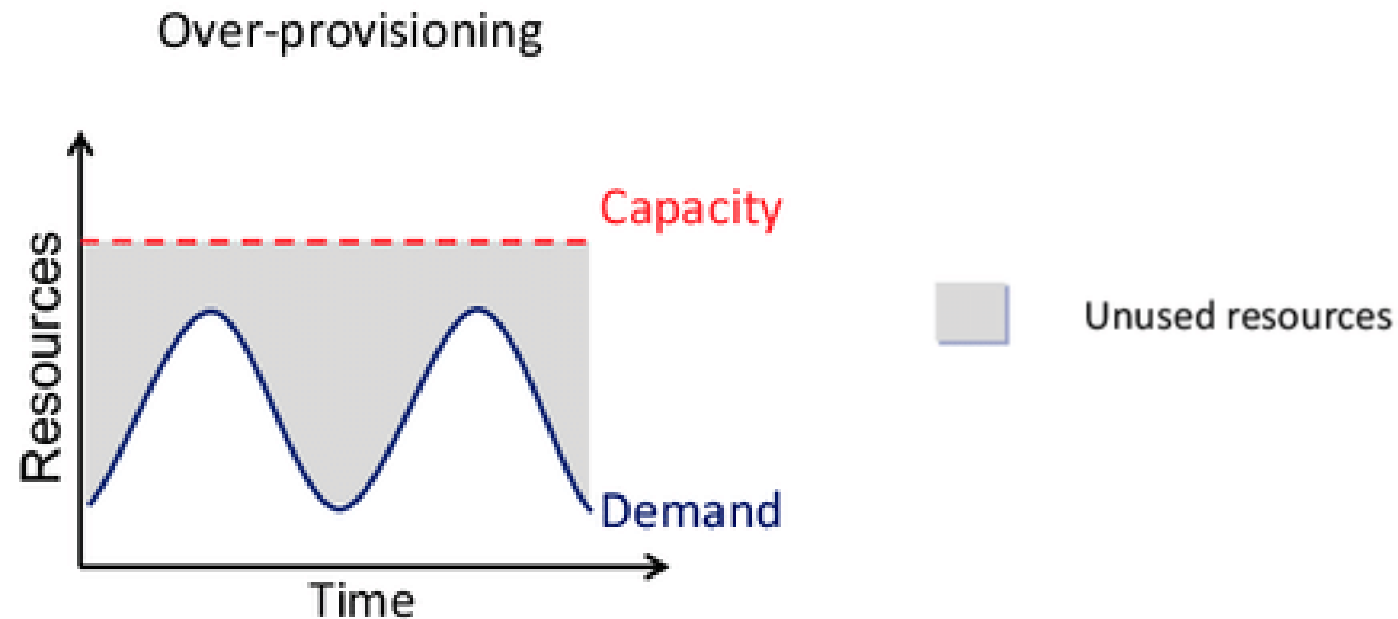


 Unused resources

- In traditional computing model, two common problems
Underestimate system utilization which result in under provision



- Overestimate system utilization which results in low utilization



- How do we solve this problem? - **Dynamically provision resources**

• There are four primary cloud deployment models

- Public Cloud
- Private Cloud
- Community Cloud
- Hybrid Cloud

• **Public clouds** are owned by cloud service providers who charge for the use of cloud resources.

- **AWS/EC2 (Amazon)**
- **Azure (Microsoft)**
- **Google Cloud Platform.**
- **Rackspace.**

• **Private clouds** are infrastructure belongs to and is operated by only one organization.

- **Eucalyptus Systems**
- **OpenNebula**
- **Openstack**

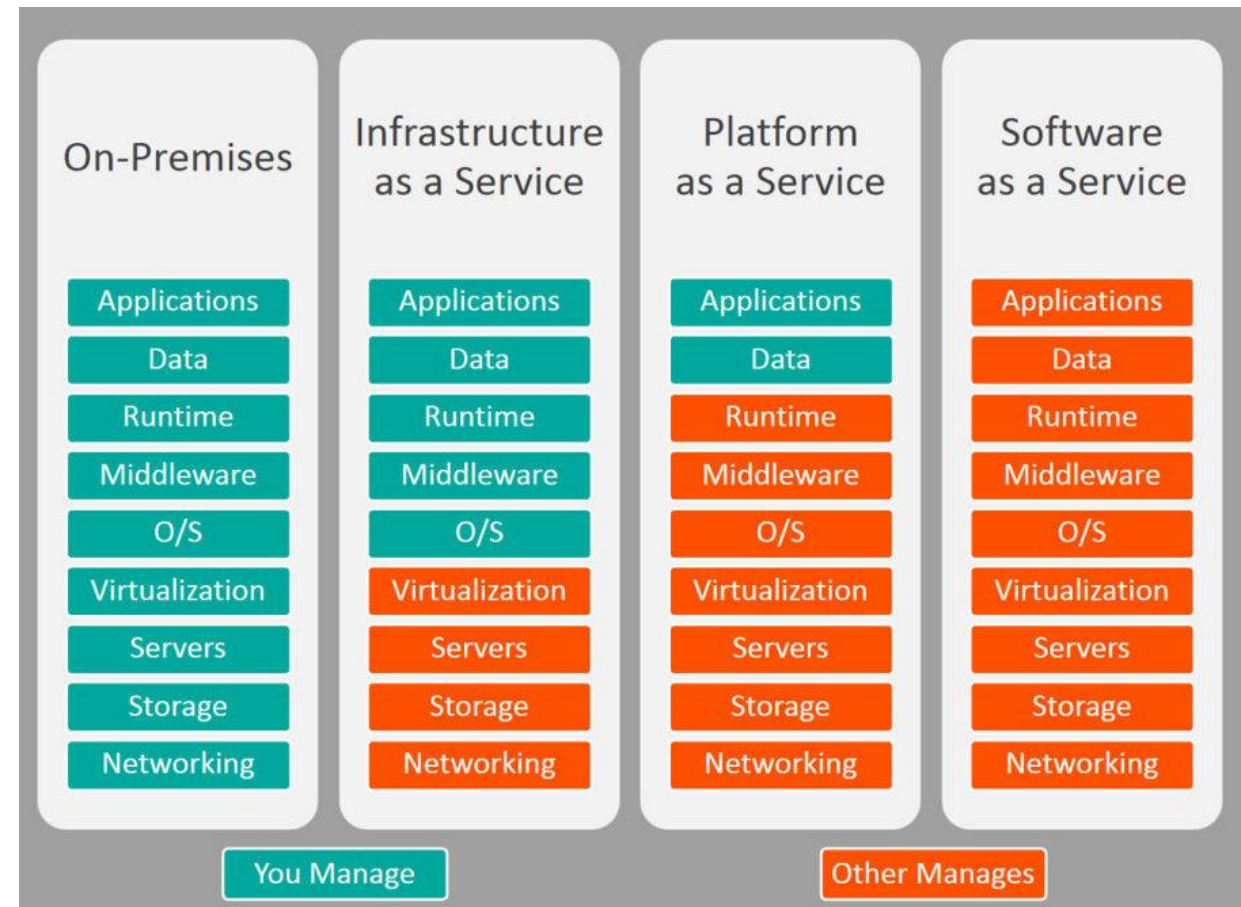
•Community cloud

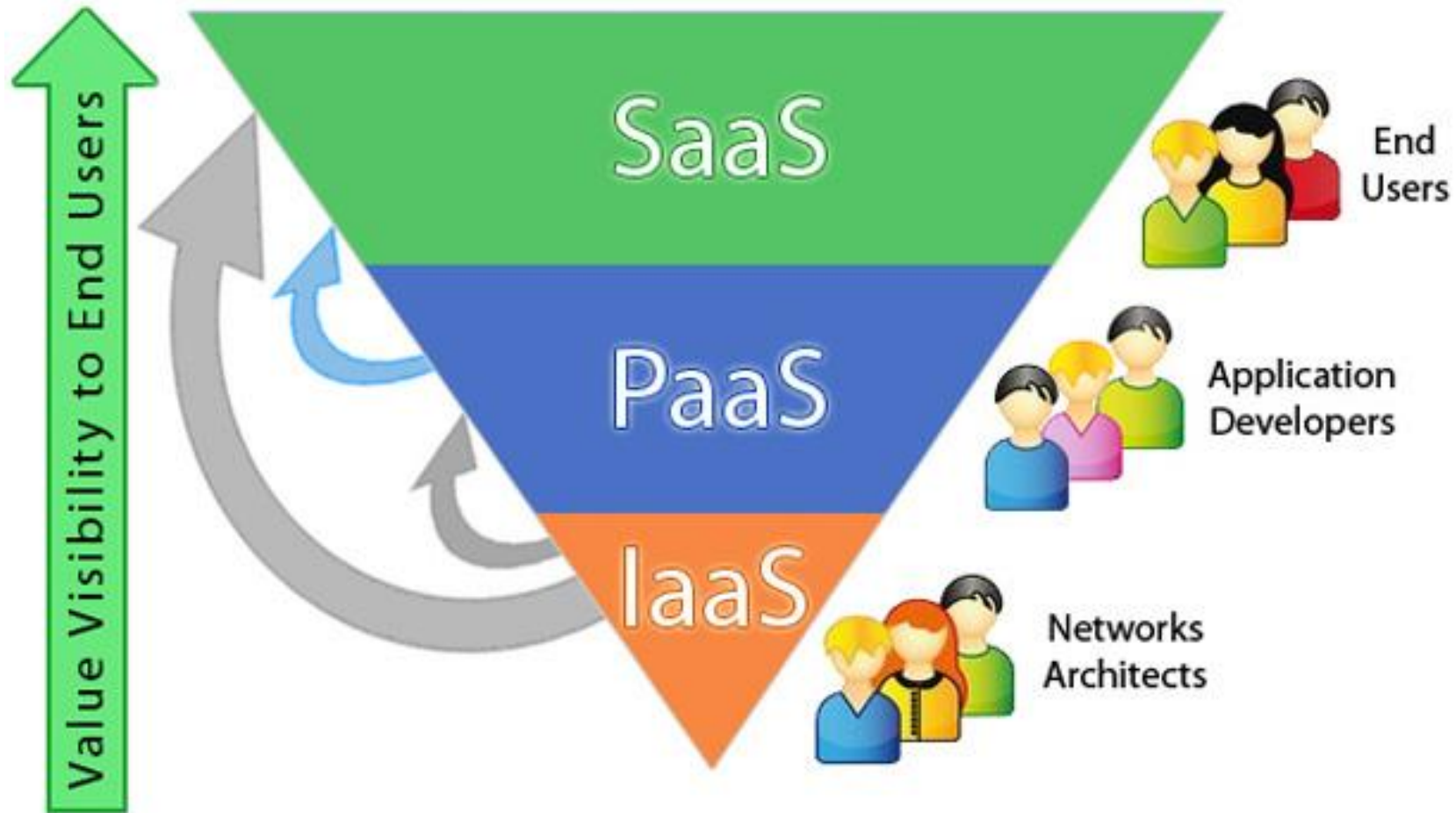
The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations).

•Hybrid cloud

The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability.

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)





- IaaS provides virtual machines, but it cannot provide elastic computing by itself, where services scale up and down to meet user demand.
 - **Dynamic provisioning**
- Existing IaaS do not provide support for the sharing middleware platforms among different VMs
 - **Multi-tenancy**

- Multi-tenancy is where a single instance of the software runs on a server, serving multiple clients.
 - Think multiple users in a MySQL database
 - Java 9 should support multi-tenancy (many java programs running in the same JVM)
- The software should be able to provide a single service to all customers by setting configurations
 - More efficient use of server resources

- **Platform as a Service (PaaS)** is a computing platform that abstracts the infrastructure, OS, and middleware to drive developer productivity.
- PaaS leverages **dynamic provisioning**
- PaaS leverages **multi-tenancy**

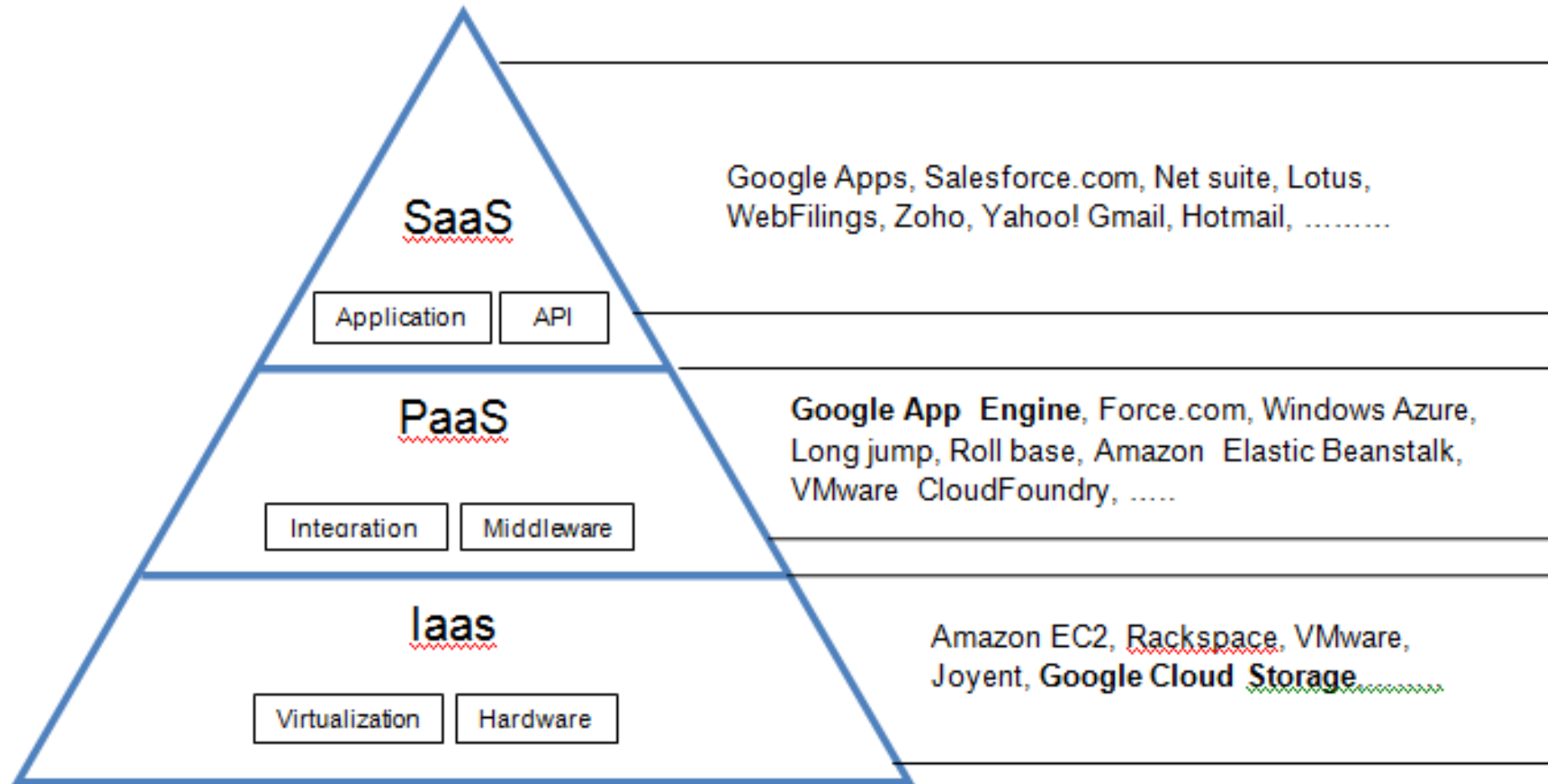
- **A closed PaaS** provides a fixed set of services you can use. You cannot install your own services.
- **An open PaaS** provides support for you to develop your own automated service deployments.

- **Software as a Service - SaaS**

- Run applications on a provider's on a cloud infrastructure.
- Applications are accessible from various client devices through a thin client interface such as a web browser.
- User is oblivious to the underlying cloud infrastructure

- **Examples - Dropbox**

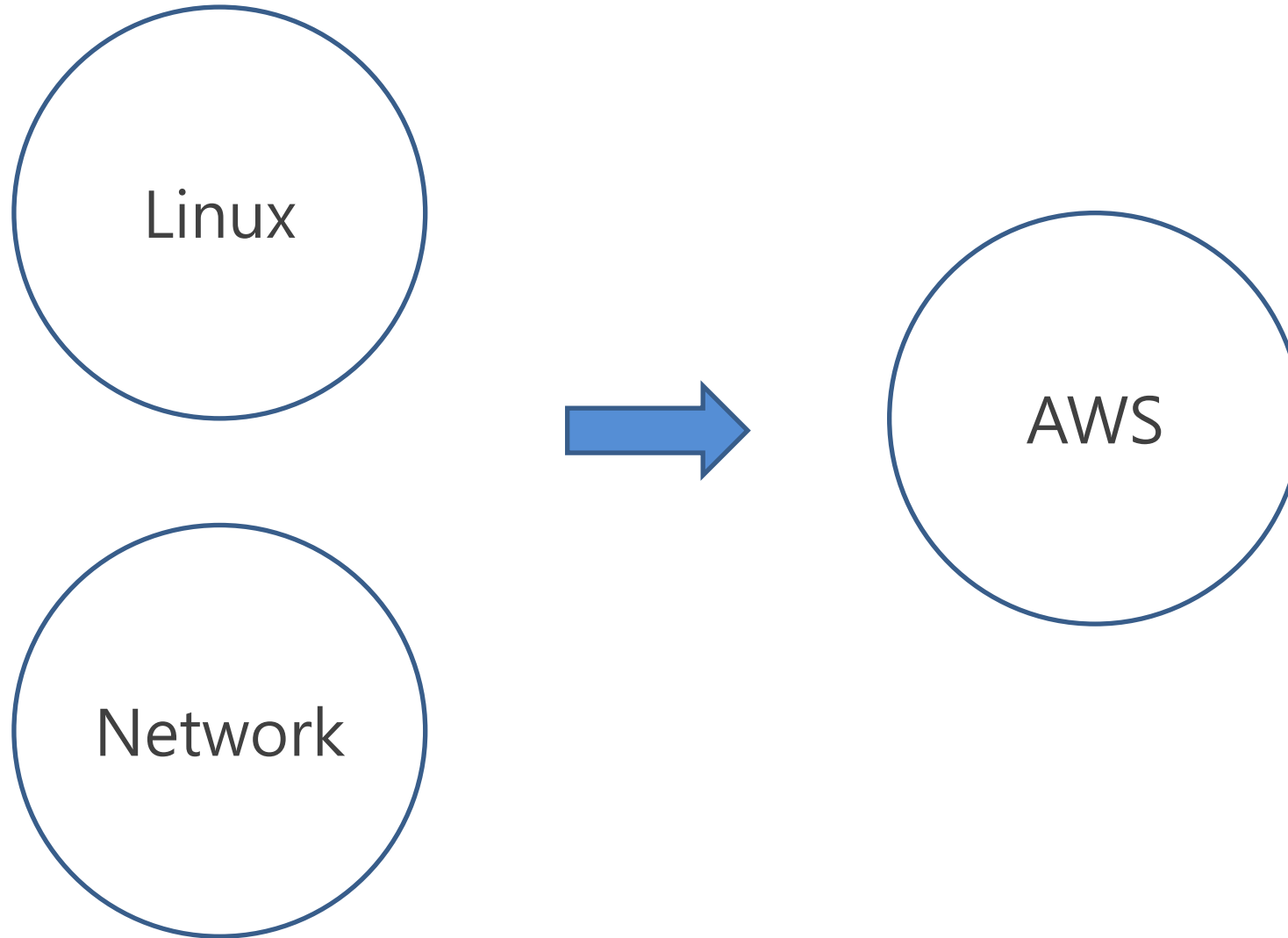
- Google Apps (e.g., Gmail, Google Docs, Google sites,..)
- Salesforce.com



- Cloud computing has enabled an explosion in largescale computing services and applications.
- Clouds provide services at three main levels: IaaS, PaaS, SaaS.
- New programming models enable easier development of large-scale applications.

2 Study Roadmap

JEONBUK NATIONAL UNIVERSITY



감사합니다

JEONBUK NATIONAL UNIVERSITY