Basics of Cloud Computing

Tessema Mengistu (Ph.D.)

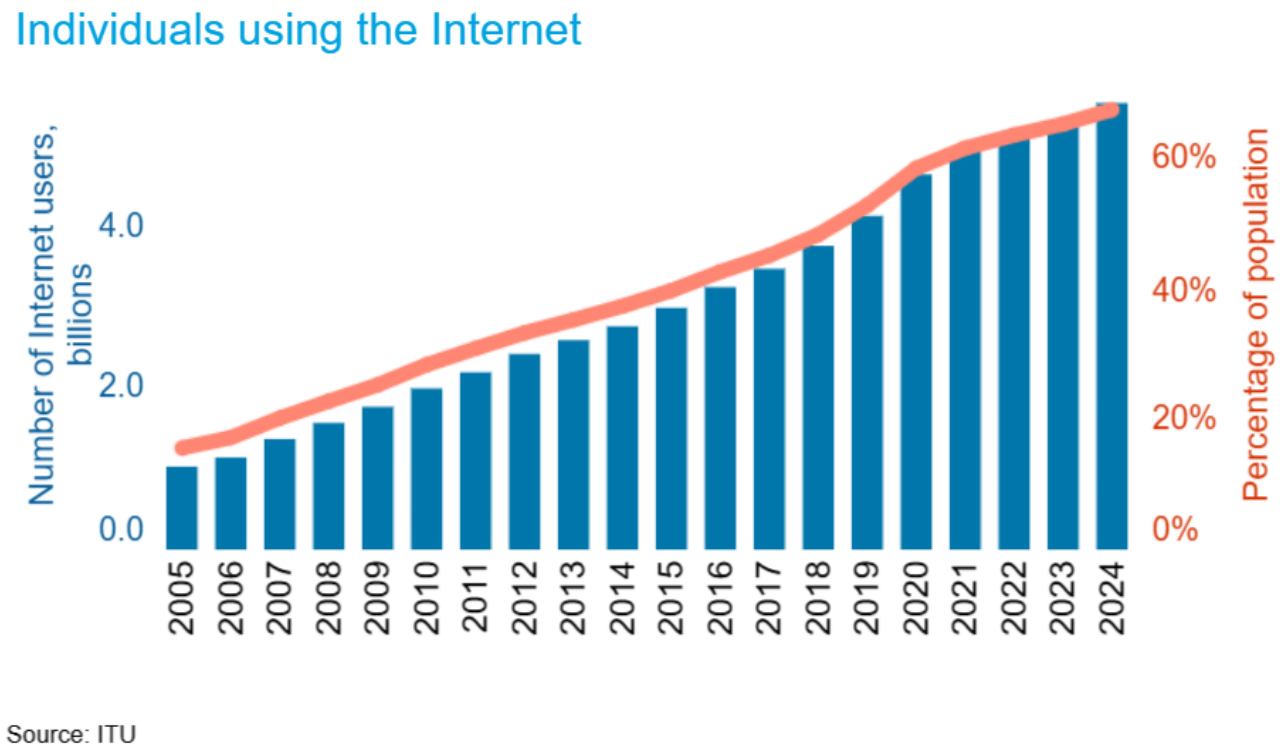
mengistu@vt.edu

Outline

- Current Trends in Computing
- Cloud Computing Introduction
  - Characteristics
  - Service Models
  - Deployment Models
- Enabling Technologies
- Terminologies

Current Trends in Computing

- The number of Internet users is increasing rapidly



Source: https://www.itu.int/en/ITU-D/Statistics/pages/stat/default.aspx

Current Trends in Computing

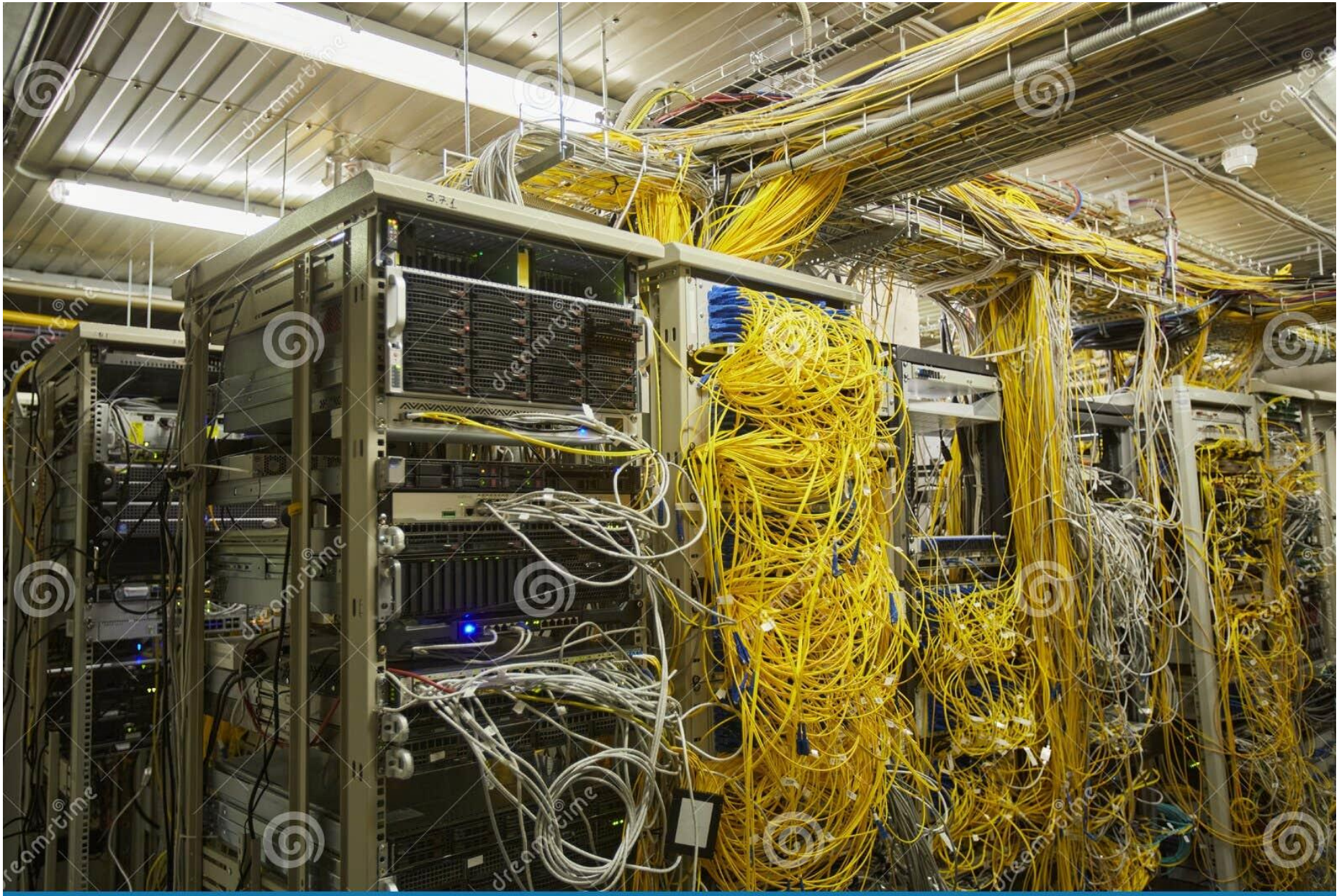
- A shift from being data consumers to data producers to intelligence
  - Web 2.0/3.0
- The explosion of devices connected to the Internet
  - Internet of Things – IoT
- Results in:
  - The need for high computational resources

Current Trends in Computing

- Infrastructure models:
  - On-premise
  - Cloud
    - Multi-cloud
  - Hybrid

Current Trends in Computing

- On-premise
  - Hardware infrastructure owned and administered by the enterprise
    - Server racks
    - Network infrastructure
    - Cooling
    - . . .
  - Advantage
    - Full control
  - Disadvantage
    - Cost
    - Limited Scalability


dreamstime.com
ID 211127506 © Seventyfourimages

Introduction to Cloud Computing

If computers of the type I have advocated become the computers of the future, then computing may someday be organized as a public utility, just as the telephone system is a public utility...

John McCarthy - 1961

Introduction to Cloud Computing

- Cloud Computing
  - is a model for enabling ubiquitous, convenient, **on-demand network** access to a **shared pool** of configurable computing resources that can be **rapidly provisioned and released** with **minimal management effort** or service provider interaction. --- NIST
- Different resources
  - Compute
  - Storage
  - Database
  - Networking
  - Applications
  - . . .

Introduction to Cloud Computing

- The **five characteristics** of Cloud Computing:
  - On-demand self-service
  - Broad network access
  - Resource pooling
  - Rapid elasticity
  - Measured service

Characteristics of Cloud

- On-demand self-service
  - A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider
- Broad network access
  - Capabilities are available over the network and accessed through standard mechanisms
- Resource pooling
  - Multi-tenant model. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter)
  - Example: Amazon Northern VA

Characteristics of Cloud

- Rapid elasticity
  - Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in.
    - Often appear to be unlimited and can be purchased in any quantity at any time
- Measured Service
  - Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts)

Characteristics of Cloud

- Additional characteristics of Cloud Computing
  - Scalability
    - Infrastructure capacity allows for traffic spikes and minimizes delays
    - Vertical and horizontal scaling
  - Resiliency
    - Mirrored solutions to minimize downtime in the event of a disaster.
    - Give businesses the sustainability they need during unanticipated events

Characteristics of Cloud

Conventional Cloud

- Cloud Computing has been catering the resource demands of most application
- "Data Center" approach.
  - Contains tens (if not hundreds) of thousands of servers.
- Utility computing vision of Cloud Computing is almost achieved.

Source: https://www.siebeswart.nl/image/I0000hjK4X4MnxPA

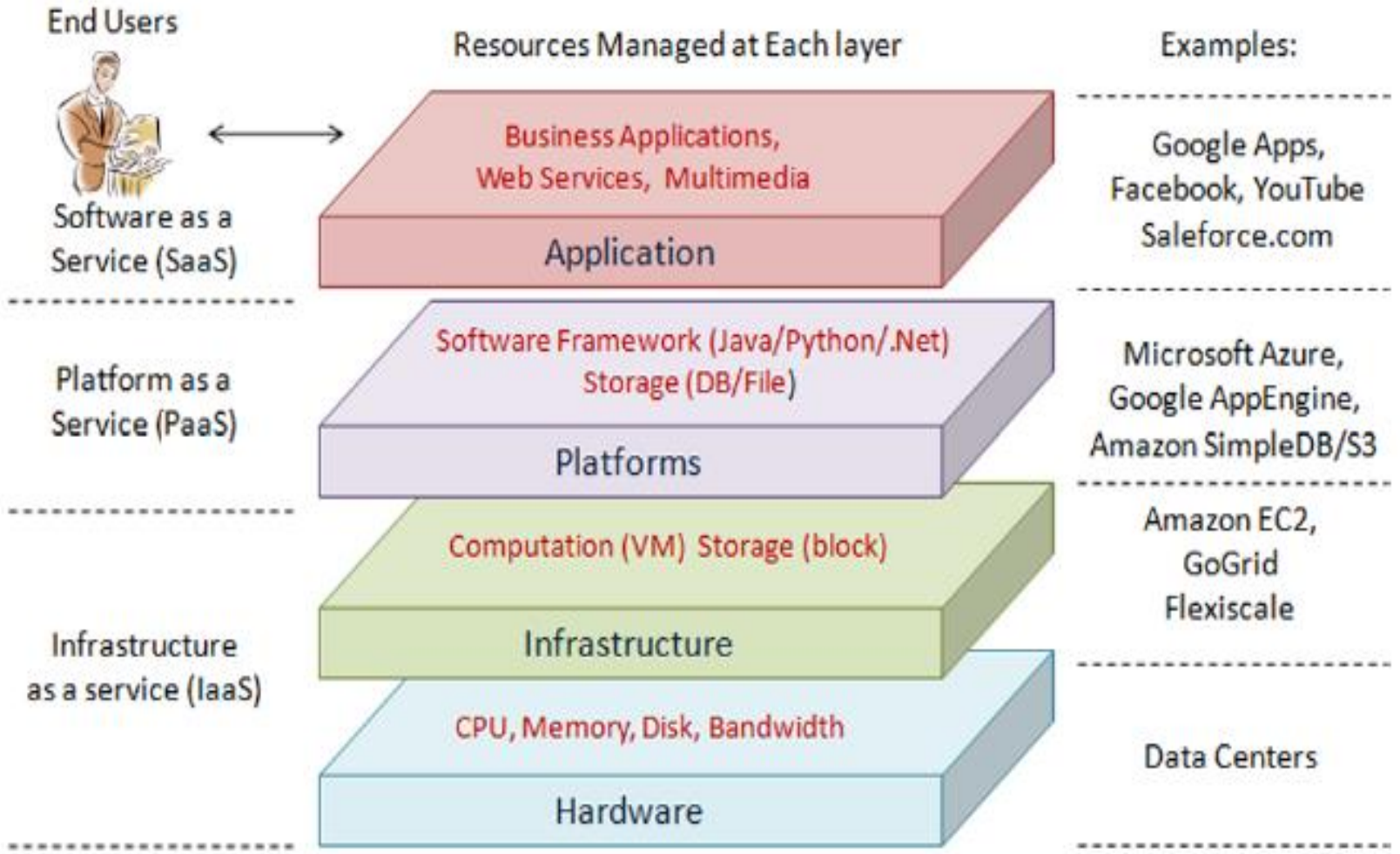
Cloud Service Models

- Cloud Computing is offered in three **basic** service models:
  - IaaS (Infrastructure as a Service)
  - PaaS (Platform as a Service)
  - SaaS (Software as a Service)

Cloud Service Models

- IaaS
  - Infrastructure as a Service
  - Delivers compute infrastructure(servers, data center space or network equipment), typically a platform virtualization environment, as a service.
- PaaS
  - Platform as a Service
  - Deliver a computing platform where the developers can develop their own applications
  - Example: Programming languages, application frameworks, databases, etc.
- SaaS
  - Software as a Service
  - A model of software deployment where the software applications are provided to the customers as a service
  - Example: Gmail, salesforce.org, etc.

Cloud Service Models



Source: unknown

Cloud Deployment Models

- Cloud computing offers typically four cloud deployment models:
  - Private clouds
    - For exclusive use by a single organization
  - Public clouds
    - Open use by the general public
  - Community Clouds
    - Shared by several organizations with shared concerns
  - Hybrid clouds
  - Distributed Clouds??
  - Industry Clouds*

https://www.gartner.com/en/articles/what-are-industry-cloud-platforms

Advantages of Cloud Computing

- Cost
  - Eliminates capital expenditure and reduce admin costs
  - Economies of scale
- Elasticity
- Infrastructure deployment speed, scalability, and agility
- Global presence

Advantages of Cloud Computing

- Ethical Issues
  - The control is relinquished to third party services
    - Security
  - Vendor Lock-in
  - Data protection and privacy

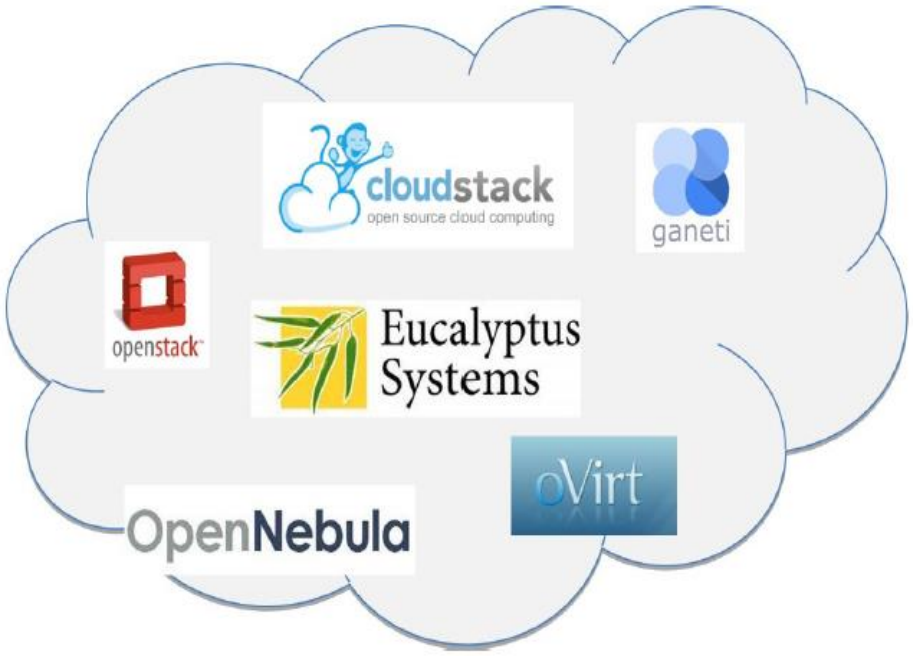
Cloud Service Providers

- Major commercial cloud providers:
  - Amazon Web Service (32%)
  - Microsoft Azure (20%)
  - Google Cloud (9%)
  - Alibaba Cloud (6%)
  - IBM Cloud

Source: https://www.zippia.com/advice/cloud-adoption-statistics/

Cloud Service Providers

- Open-source cloud platforms:



Cloud Computing Enabling Technologies

- Enabling technologies
  - Service Oriented Architecture (SOA)
  - Virtualization
  - Multicore technology
  - Datacenter technology
  - Networking

Terminologies

- A Service Level Agreement (SLA)
  - A negotiated contract between the customer and the service provider
  - The objectives of the agreement are:
    - Identify and define the customer's needs and constraints including the level of resources, security, timing, and quality of service
    - Provide a framework for understanding; a critical aspect of this framework is a clear definition of classes of service and the costs
    - Simplify complex issues; for example, clarify the boundaries between the responsibilities of the clients and those of the provider of service in case of failures.
    - Reduce areas of conflict
    - Eliminate unrealistic expectations

Terminologies

- An SLA records a common understanding in several areas:
  - Services
  - Metrices
  - Responsibilities
  - Guarantees
  - Warranties

References

- A View of Cloud Computing [ACM, 2010]
- Cloud Basics [Google]
- NIST
  - The NIST Definition of Cloud Computing [NIST]
  - NIST Cloud Computing Reference Architecture [NIST]