

Cloud Computing Fundamentals

Week One

School of Computer Science

Faculty of Engineering and Information Technology

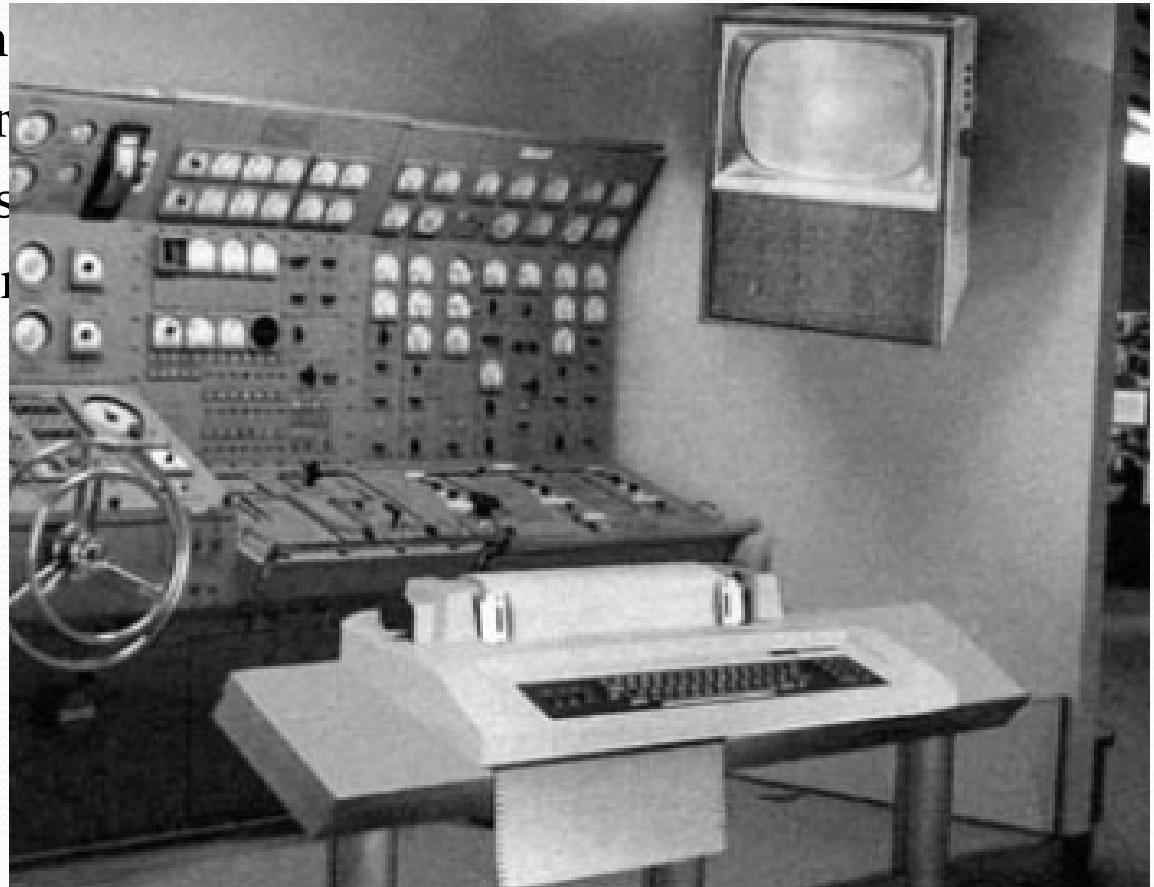
University of Technology Sydney

Learning Objectives

- Overview of the evolution of computing
- Understand the emergence of cloud computing
- Definition of cloud computing
- Understand the key parameters of cloud computing
- Understand cloud computing driving parameters

Evolution of Computing Hardware (1/3)

- Single stand-alone machine (1950's)
 - Big Bulky machine
 - Limited Memory
 - Limited Processors
 - Extremely expensive



"I think that there is a world market for maybe five computers" -- Thomas Watson, Chairman of IBM, 1943

Evolution of Computing Hardware (2/3)

- Commercial Mainframes (1960's)
 - Big “Frames” comprising of computing elements
 - Initially were developed for business
 - Time-Shared (virtual machines)
 - Typical memory was 1-2 million words
 - Typical processing power was 1-2 million instructions per second



"Where a calculator on the ENIAC is equipped with 18,000 vacuum tubes and weighs 30 tons, computers in the future may have only 1,000 vacuum tubes and weigh only 1.5 tons." -- *Popular Mechanics*, 1949

Evolution of Computing Hardware (3/3)

- Personal Computers (1980's)
 - Pioneered by IBM
 - Advancement in Microprocessors
 - Typical memory: 16 KB
(256 KB in the powerful versions)
 - Typical Non-Volatile memory: 4 MB
 - Typical Processing Power 4 MHz – 8 MHz

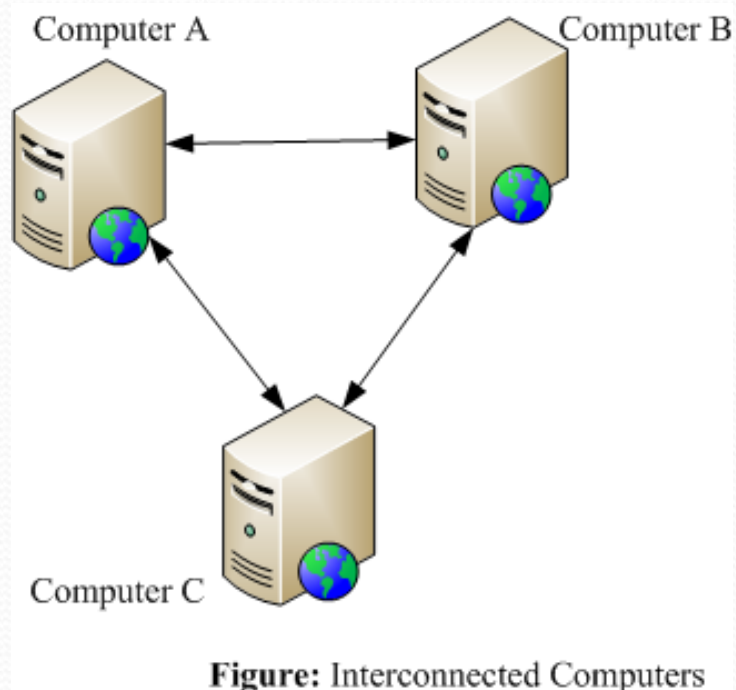


Figure : An IBM PC

"640KB ought to be enough for anybody." -- Attributed to Bill Gates, 1981, but believed to be an urban legend

World Wide Web

- Implementing of the World Wide Web (1990's)
 - Building a network of computers
 - Standardization of networking technology (TCP/IP)
- Large scale adoption of the World Wide Web in 1990's



Grid Computing (Late 1990's)

- Utilization of idle computing resources
- “*Knitting*” of autonomous computing resources
- Harnessing the power of the web for solving complex problems that cannot be solved by a single computer
- Geared towards scientific problem solving (geological sciences, medical sciences etc....)

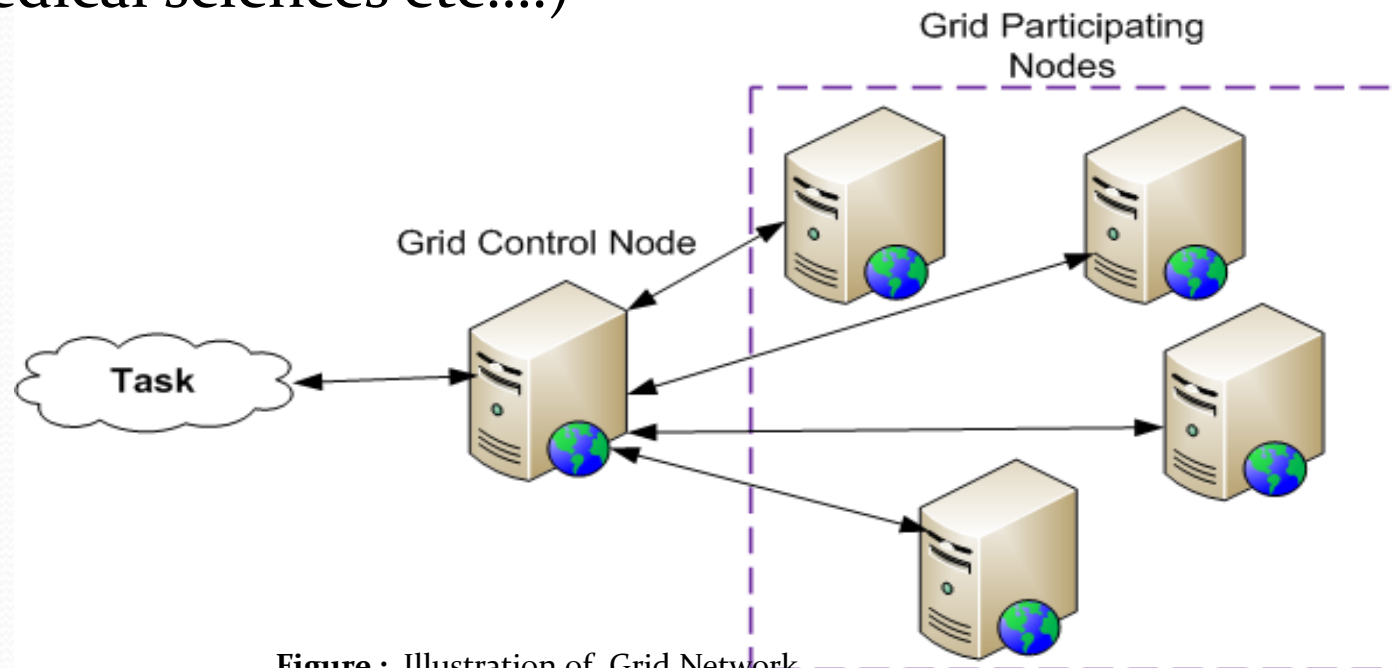


Figure : Illustration of Grid Network

Evolution in Computing Software

- Changes in the design philosophy for engineering computing software
- Process-Oriented Architecture
 - Programs are functionally decomposed into processes, that act upon shared data
 - Data structures are complex and large

Object-Oriented Architecture

- Reflection of the real world comprising of objects or entities
- An object is an encapsulation of a data structure, and the methods or operations to access those structures
- Objects communicate with each other by passing messages (which in turn may invoke other object)

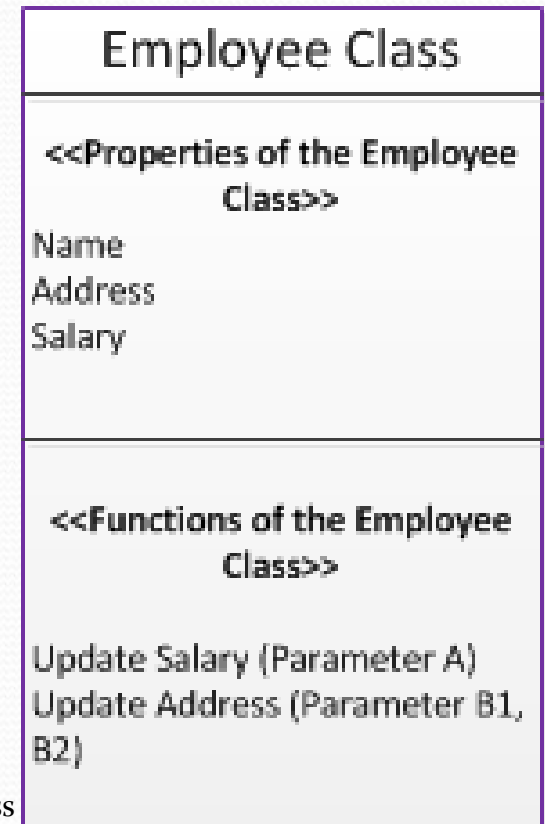


Figure : An example of class

Service Oriented Architectures

- The software application is construed as being organized of “*software services (components)*”.
- A “software service” is a piece of code that provides a given (well defined) functionality (Example: “Authentication”, “Registration”, “Billing” etc...).
- A software service can be “*invoked*” to deliver its functionality
- Software service interfaces are based on public standards for interoperability.
- Promotes re-use of software services across a range of applications (Examples: Registration Component, Authentication Component, Billing Component etc... , and their use across a number of software applications such as Student Management Application, Staff Management Application, Bank Applications etc...)

Service Oriented Architectures

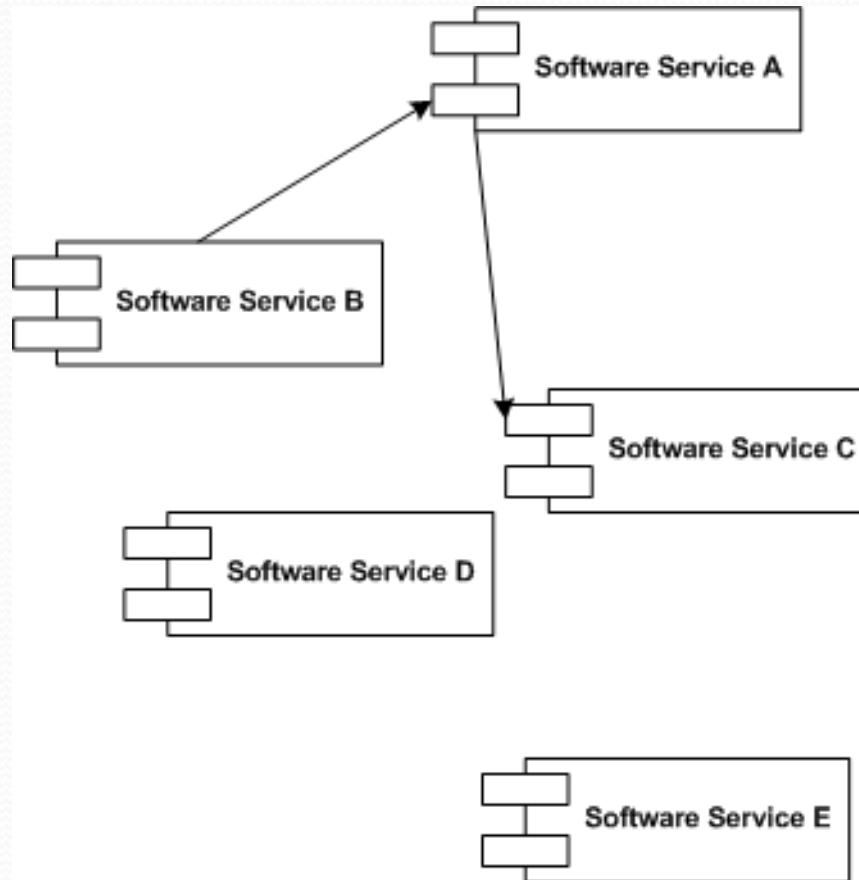
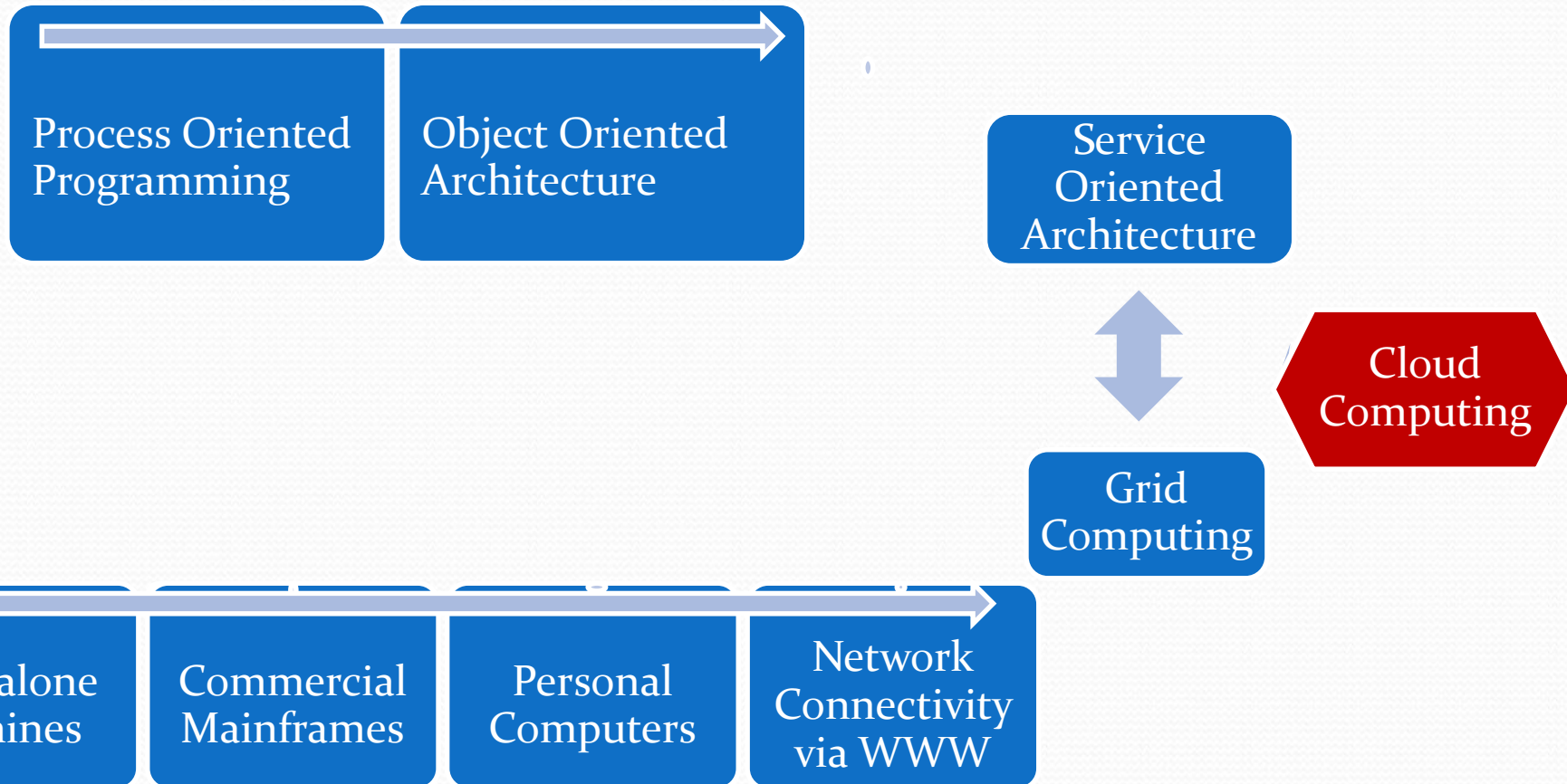


Figure : Illustration of software services invoking

Evolution of Cloud Computing

An Integrated View



Cloud Computing

The NIST Definition of Cloud Computing

“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”.

Key aspects of the definition:

- (a) Cloud providers own a large pool of computer resources
- (b) Computing resources are offered “as-a-services”
- (c) Computing resources can be dynamically scaled-up or vice-versa

Key parameters of Cloud Computing

- (a) Pooled (computing) resources
- (b) Elastic scaling of resources
- (c) Virtualization of resources
- (d) Pay-as-you-go (or) Metered resource usage
- (e) Resource Delivery over the internet
- (f) Automation of new resource deployment and de-provisioning

Pooling of Computing Resources

- Computing Resources (are pooled) by the (cloud) service provider, and made available to the requestors (cloud service requestors).
- Available on request to anyone without any (geographical) constraints

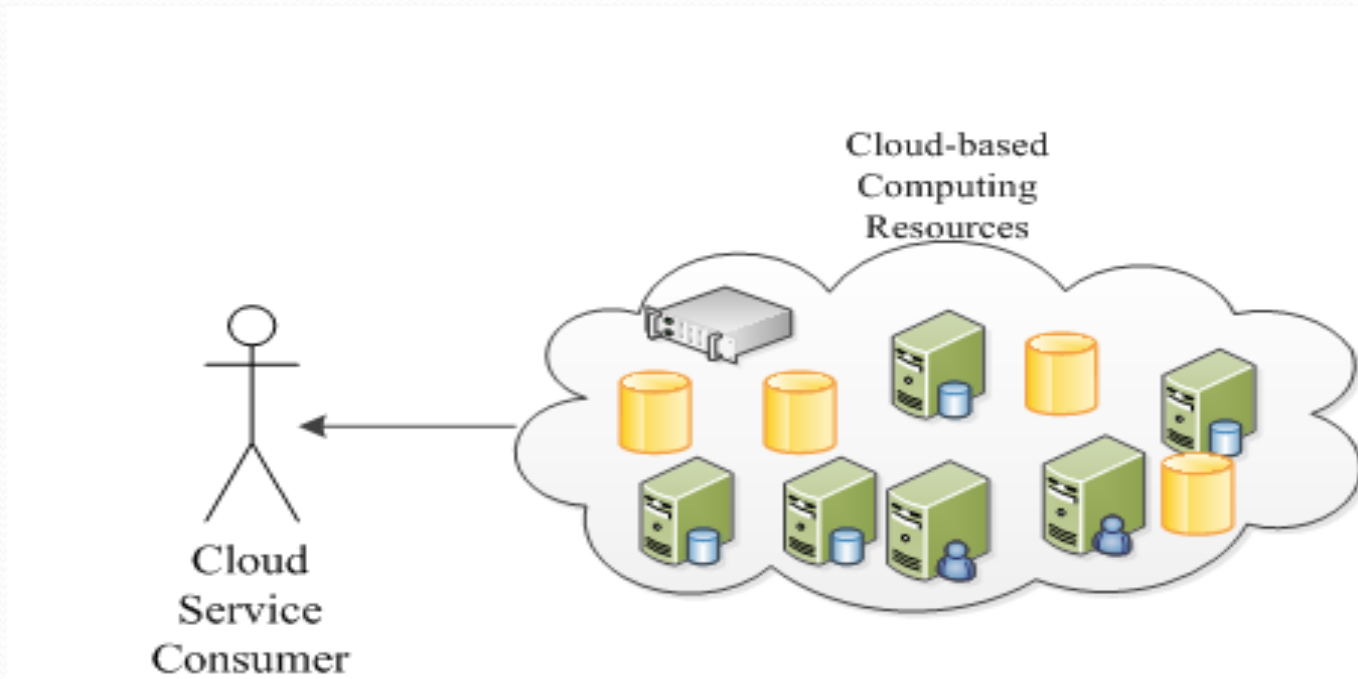


Figure : Illustration of pooling of computing resources

Elastic scaling of computing resources

- Applications require base level computing resources under normal circumstances
- They may require additional resources to handle abnormal circumstances (peak load)
- Elastic scaling of computing resources
 - a) Cloud platforms offer dynamic scale-up or scale-down of computing resources
 - b) Saves resources and time to handle peak demand
- What would be economic implications of (a) and (b) above for enterprises?

Elastic scaling of computing resources

- Move from investing in Capital Expenses (CapEx) to Operational Expenses (OpEx)
 - No-upfront costs
 - Reduction in the actual value of Operational Expenses due to the economies of scale.

Virtualized computing resources

- Enables high utilization of the computing resources
- “*Slice-and-dice*” computing resources to create virtual resources
 - Virtual resources are provided to cloud consumer (on-the-fly)
- Virtualization is enabled through virtualization mechanisms such as hypervisors etc...
- Example (Virtual Servers):
 - Physical Servers can be partitioned to a number of “virtual servers”
 - Virtualization mechanisms ensure that each virtual server acts like a real server

Virtualized computing resources

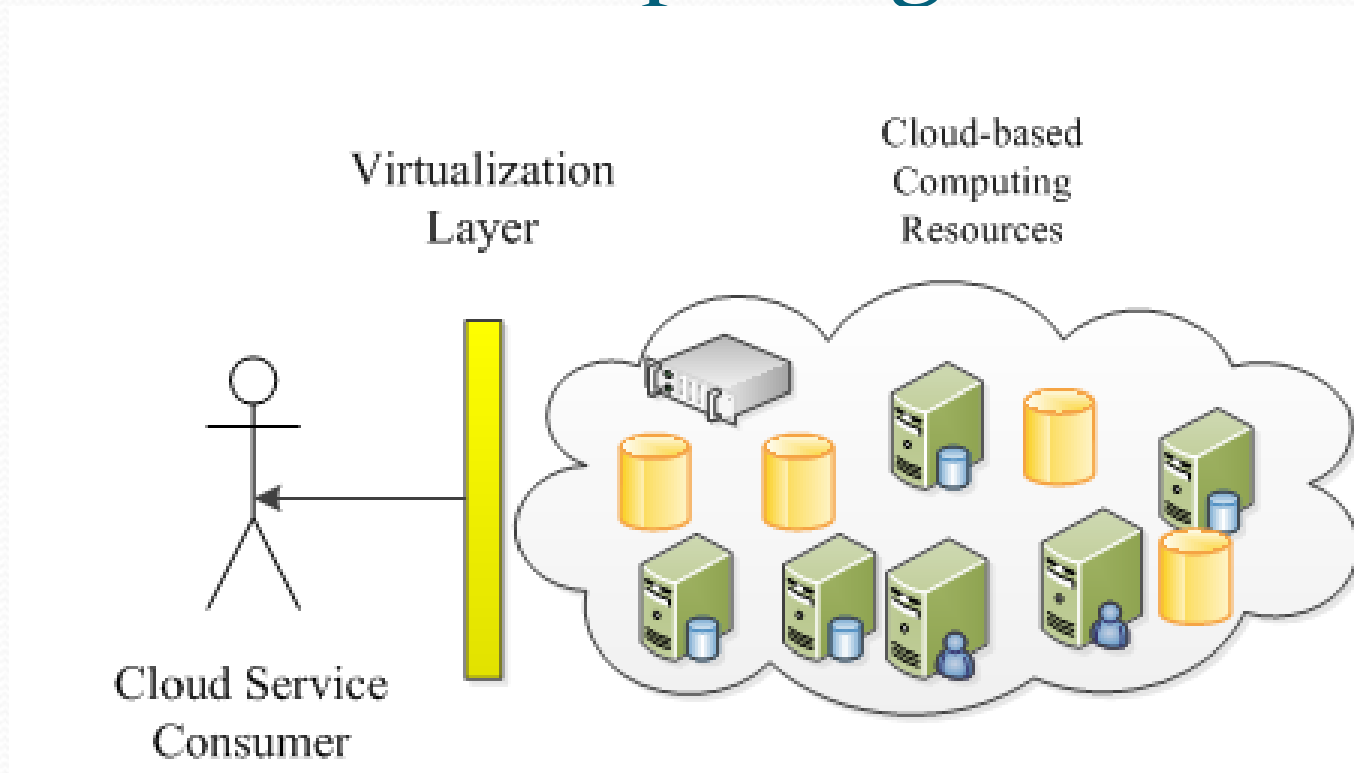


Figure : Illustration of relationship between cloud resources, virtualization and cloud consumer

Metered computing resource usage

- Typically no start-up fees or set-up fees
 - Sign up, use cloud resource, pay for your usage (based on your consumption of resources)
 - Enables SME's to use large-scale IT resources (by removing the resource procurement barrier)
 - Enables Large Enterprises to consume resources cheaply
 - Fine-Grained Metering and Billing of Resource Usage



Figure : An example of Cloud service pricing

Resource Delivery over the internet

- (Computing) resources are provisioned as a “*service*”
- Resources are delivered over the internet
- Typical services include (hardware (infrastructure); computing platforms; and software)
- API's are provided to access and use the provided resources

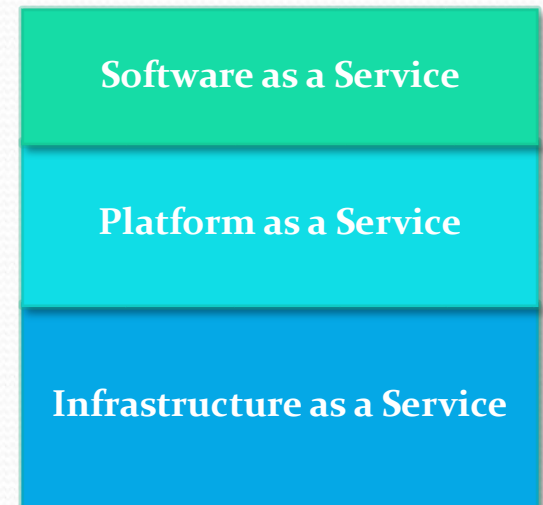


Figure : Different types of Cloud services

Resource Deployment and De-provisioning Automation

- New or additional resources are requested via an API
- The requested resources by the cloud consumer will be configured and made ready to use (a.k.a Resource Provisioning or Deployment)
- The provisioned resources will be de-provisioned when they are not longer required
- No effort or action (apart from informing) from the cloud consumer to provision, configure, maintain, update or de-provision the resources
- Saves time and resources in configuring and deploying of resources

Benefits of Cloud Computing

- Reduction in business costs
 - CapEx to OpEx (Quantum Shift in capital requirements)
 - Cheaper OpEx costs
 - Reduction of space usage
 - Huge reduction in time and man-power required for planning, purchasing, provisioning, configuring, and disposing resources
- Technology to Core Business Shift
 - Enables organize to focus on core-business activities
 - Gain competitive advantage by focusing on core-business activities



History of Amazon Cloud

- Amazon Cloud

- Started as an on-line book store, and later expanded its product range in the e-Commerce market
- Huge growth in online-presence over the years
- Innovative business strategies (such as Amazon Merchants and Market places) exacerbated business growth
- Business growth implied large data storage and computing resources from Amazon
- Peak sales duration around December
- Led to overprovisioning of computing resources during the rest of the year
- Amazon Web Service launched in 2002 sold idle computing resources to other organizations who needed them

Utility computing

- The notion that computing resources can be organized and distributed as a utility (similar to water, electricity, gas, and internet)
- Computing as a “fifth utility” for human beings
- Key aspects of utility computing:
 - Computing resources can be requested and provided to any one on demand;
 - Seamless up-scale and down-scaling of resources from the utility providers;
 - Low or no initial set-up costs;
 - Pay-as-you go resource usage
- Cloud computing is one of the mechanisms for achieving utility computing

(Some) Cloud Computing Applications

- Google Products such as
 - Google Mail (GMail)
 - Google Calendar
 - Google Docs
 - Google Apps
- Microsoft Azure
- Dropbox
-
-
- Amazon Products such as
 - Amazon Elastic Cloud Compute (EC2)
 - Amazon Storage (S3 etc..)

Reading

Books

1. Rhoton, J. (2010), Cloud computing explained, Recursive Press, UK – Chapters 1, 3-6
2. Shroff, G. (2010), Enterprise cloud computing: technology, architecture, application, Cambridge University Press, UK – Chapters 1 and 2
3. Moyer, C.M. (2011), Building Applications in the Cloud: Concepts, Patterns and Projects, Addison-Wesley, Pearson Education – Chapter 1

Reading

Papers, Websites

1. P. Mell and T. Grance. (2009), The NIST Definition of Cloud Computing.
<http://www.nist.gov/itl/cloud/upload/cloud-def-v15.pdf>.
2. M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. (2009). Above the clouds: A Berkeley view of Cloud Computing, UC Berkeley EECS.
3. Open Cloud Manifesto (2009),
<http://www.opencloudmanifesto.org/>