Introduction to Cloud Computing

CSE-6305 Cloud Computing

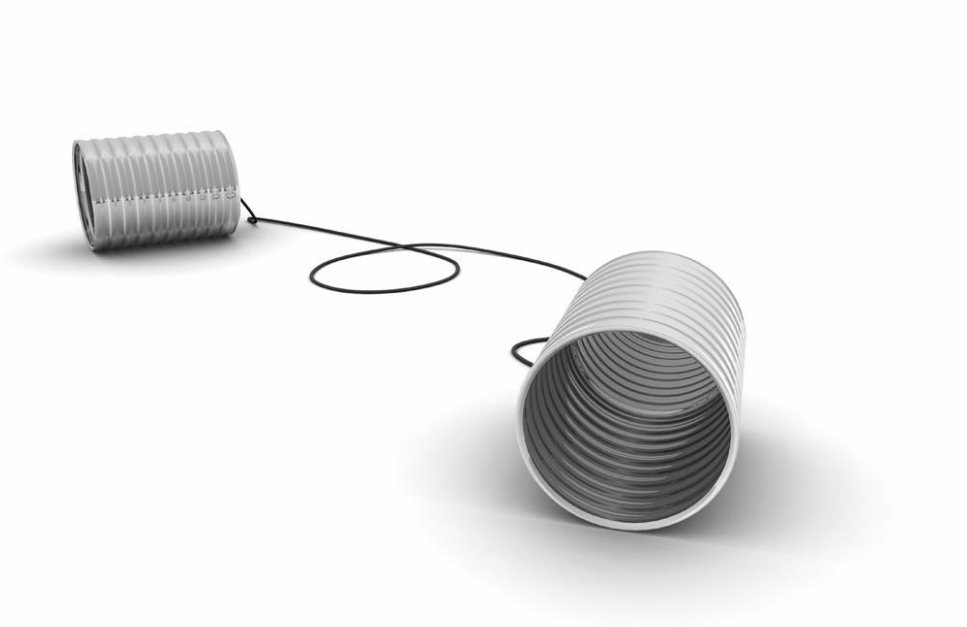
Professor Dr. Mohammad Abdur Rouf

Class Administration

- Lecture Material is available:
  - https://sites.google.com/a/duet.ac.bd/marouf-cse/courses-2018/cse-6305-cloud-computing
- Marks Distribution
  - Attendance: 10%
  - Midterm/Homework: 20%
  - Final Exam: 70%
- Text Book:
  - Will be provided during lecture
- Email: marouf.cse@duet.ac.bd, rouf7606@gmail.com

Agenda

- What is Cloud Computing ?
  - Different perspectives
  - Properties and characteristics
  - Benefits from cloud computing
- Service and deployment models
  - Three service models
  - Four deployment models



What is Cloud Computing ?

What do they say ?

Cloud Disclaimers

- Talk from Oracle CEO ***Larry Ellison***
  - We've redefined Cloud Computing to include everything that we already do. I don't understand what we would do differently other than change the wording of some of our ads.
- Talk from ***Rich Stallman***
  - It's stupidity. It's worse than stupidity: it's a marketing hype campaign. Somebody is saying this is inevitable – and whenever you hear somebody saying that, it's very likely to be a set of businesses campaigning to make it true.

Cloud Definitions

- Definition from ***NIST*** ***(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.
  - This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

- Definition from ***NIST*** ***(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.
  - This cloud model promotes **availability** and is composed of five essential characteristics, three service models, and four deployment models.

Cloud Definitions

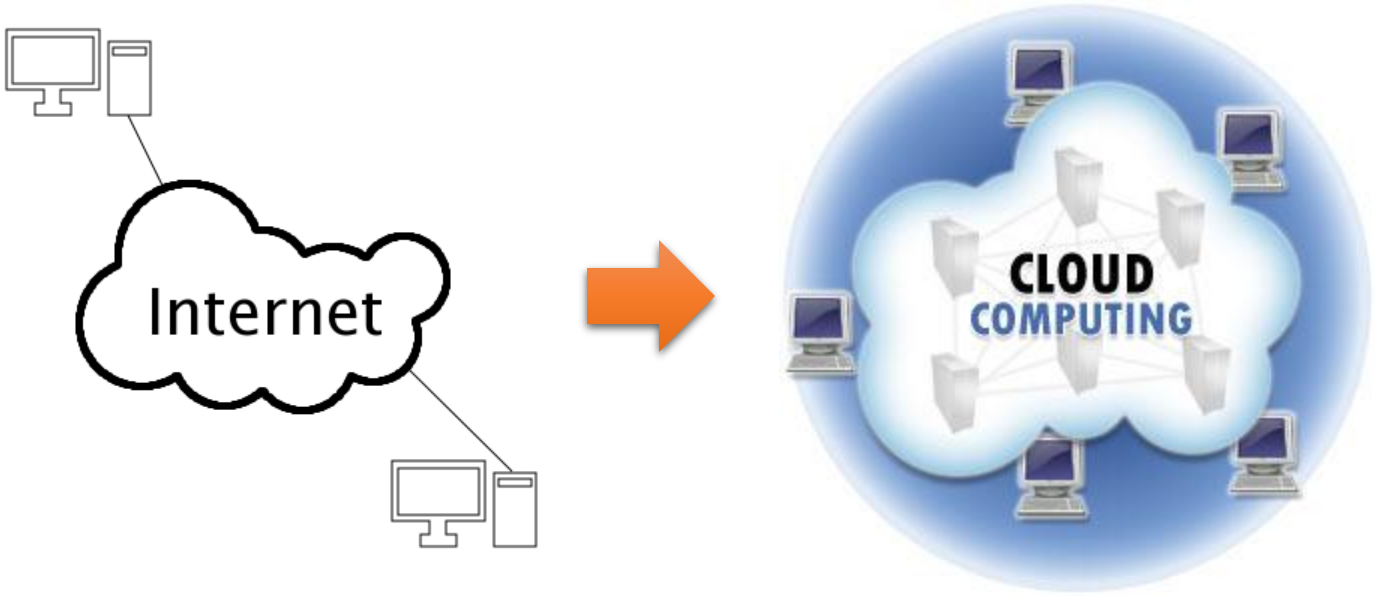
- Definition from ***Wikipedia***
  - Cloud computing is Internet-based computing, whereby shared resources, software, and information are provided to computers and other devices on demand, like the electricity grid.
  - Cloud computing is a style of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet.

WIKIPEDIA
The Free Encyclopedia

- Definition from ***Wikipedia***
  - Cloud computing is **Internet-based computing**, whereby **shared resources, software, and information** are provided to computers and other devices **on demand**, like the electricity grid.
  - Cloud computing is a style of computing in which **dynamically scalable** and often **virtualized resources** are provided as a **service** over the Internet.

Cloud Definitions

- Definition from ***Whatis.com***
  - The name cloud computing was inspired by the cloud symbol that's often used to represent the Internet in flowcharts and diagrams. Cloud computing is a general term for anything that involves **delivering hosted services over the Internet**.



Cloud Definitions

- Definition from ***Berkeley***
  - Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services.
  - The services themselves have long been referred to as Software as a Service (SaaS), so we use that term. The datacenter hardware and software is what we will call a Cloud.
  - When a Cloud is made available in a pay-as-you-go manner to the public…… The service being sold is Utility Computing.

- Definition from ***Berkeley***
  - Cloud Computing refers to both the applications **delivered as services over the Internet** and the hardware and systems software in the datacenters that provide those services.
  - The services themselves have long been referred to as **Software as a Service (SaaS)**, so we use that term. The datacenter hardware and software is what we will call a Cloud.
  - When a Cloud is made available in a **pay-as-you-go** manner to the public…… The service being sold is **Utility Computing**.

Cloud Definitions

- Definition from ***Buyya***
  - A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers.

- Definition from ***Buyya***
  - A Cloud is a type of **parallel and distributed system** consisting of a collection of interconnected and **virtualized computers** that are **dynamically provisioned** and presented as one or more unified computing resources based on **service-level agreements** established through negotiation between the service provider and consumers.

In Our Humble Opinion

- Cloud computing is a paradigm of computing, a new way of thinking about IT industry but not any specific technology.
  - Central ideas
    - ***Utility Computing***
    - ***SOA*** - Service Oriented Architecture
    - ***SLA*** - Service Level Agreement
  - Properties and characteristics
    - High ***scalability*** and ***elasticity***
    - High ***availability*** and ***reliability***
    - High ***manageability*** and ***interoperability***
    - High ***accessibility*** and ***portability***
    - High ***performance*** and ***optimization***
  - Enabling techniques
    - Hardware virtualization
    - Parallelized and distributed computing
    - Web service

Properties and Characteristics

Central Ideas

DON'T TELL ME DETAILS!!

I DON'T CARE!!

Central Ideas

- Perspective from user :
  - Users do not care about how the works are done
    - Instead, they only concern about what they can get
  - Users do not care about what the provider actually did
    - Instead, they only concern about their quality of service
  - Users do not want to own the physical infrastructure
    - Instead, they only want to pay as many as they used
- What dose user really care ?
  - They only care about their “Service”

Utility Computing

- One service provisioning model
  - Service provider makes computing resources and infrastructure management available to the customer as needed, and charges them for specific usage rather than a flat rate.
  - Like other types of on-demand computing , the utility model seeks to maximize the efficient use of resources and/or minimize associated costs.

What Is Service?

- Service is what you connect together using Web Services.
- Service is the endpoint of a connection.
- Functionalities of service :
  - A service should be well-defined
  - A service should be self-contained
  - A service should not depend on the context or state of other services.

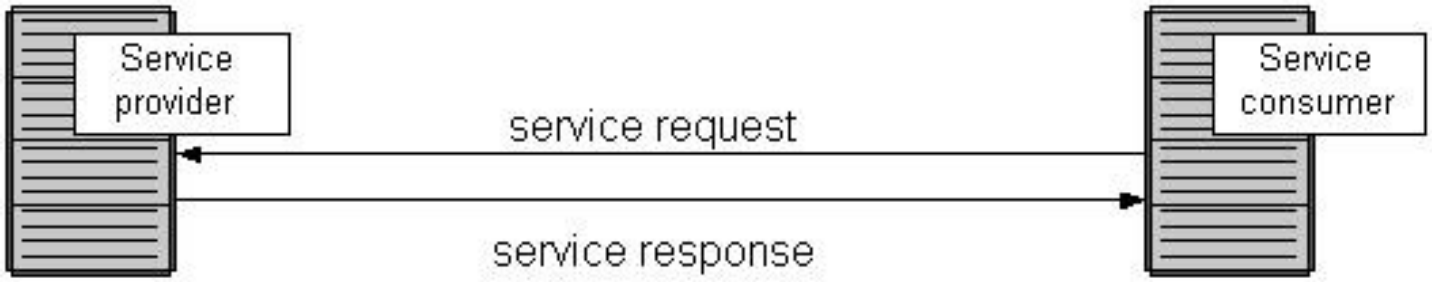
What Is Web Service?

- Definition :
  - Web service is self-describing and stateless modules that perform discrete units of work and are available over the network
  - Web service providers offer APIs that enable developers to exploit functionality over the Internet, rather than delivering full-blown applications

- Web Services Description Language (WSDL) :
  - Expressed in XML which include both data type and messages
  - Four types of operations :
    - One-way - Messages sent without a reply required
    - Request & response - Sending and replying messages
    - Solicit response - A request for a response
    - Notification - Messages sent to multiple receivers

Service Oriented Architecture

- Definition
  - Service Oriented Architecture (SOA) is essentially a collection of services which communicate with each other
  - Contain a flexible set of design principles used during the phases of systems development and integration
  - Provide a loosely-integrated suite of services that can be used within multiple business domains
- Approach
  - Usually implemented by Web Service model



Quality Of Service

- Original definition
  - Quality of Service (QoS) is a set of technologies for managing network traffic in a cost effective manner to enhance user experiences for home and enterprise environments.

- Now QoS becomes to a broad term that is used following areas :
  - Customer care evaluations
  - Technological evaluations

Quality Of Service

- Customer care evaluations
  - QoS is usually measured in terms of issues that have a direct impact on the experience of the customer
  - Only issues that produce a negative effect on the goods and services received by the customer come under scrutiny
- Technological evaluations
  - QoS has to do with the efficient operation of various systems
  - This can lead to adjusting procedures or adapting software programs and code to achieve the desired effect while making a more efficient use of available resources

Service Level Agreement

- Definition
  - A service-level agreement (SLA) is a contract between a network service provider and a customer that specifies, usually in measurable terms (QoS), what services the network service provider will furnish

- Common content in contract
  - Performance guarantee metrics
    - Up-time and down-time ratio
    - System throughput
    - Response time
  - Problem management detail
  - Penalties for non-performance
  - Documented security capabilities

Scalability & Elasticity

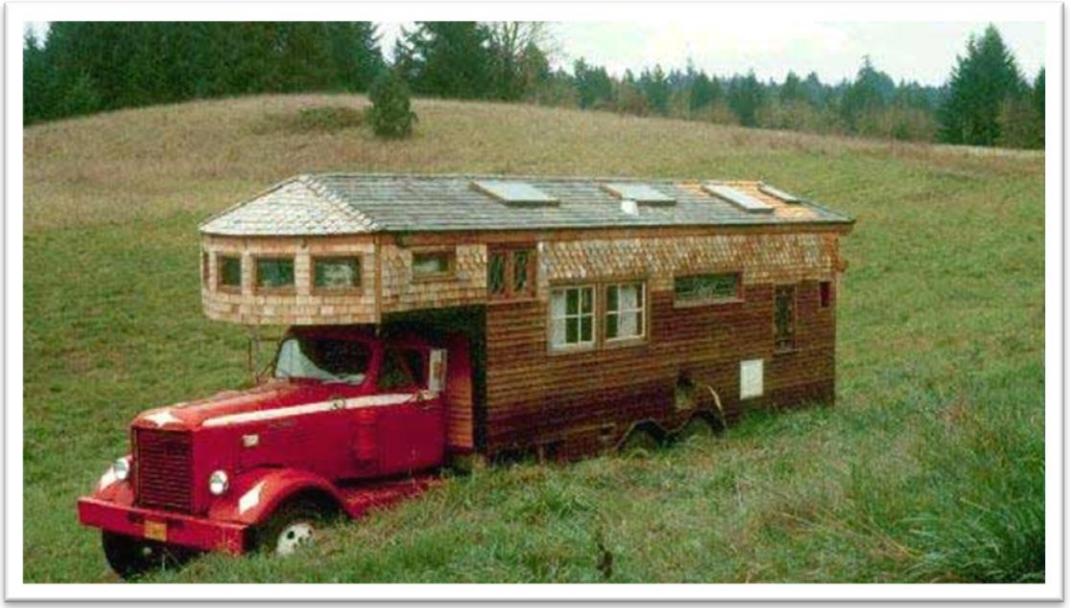
Give me the world without limitation!!

Scalability & Elasticity

- What is scalability ?
  - A desirable property of a system, a network, or a process, which indicates its ability to either handle growing amounts of work in a graceful manner or to be readily enlarged.
- What is elasticity ?
  - The ability to apply a quantifiable methodology that allows for the basis of an adaptive introspection with in a real time infrastructure.
- But how to achieve these properties ?
  - Dynamic provisioning
  - Multi-tenant design

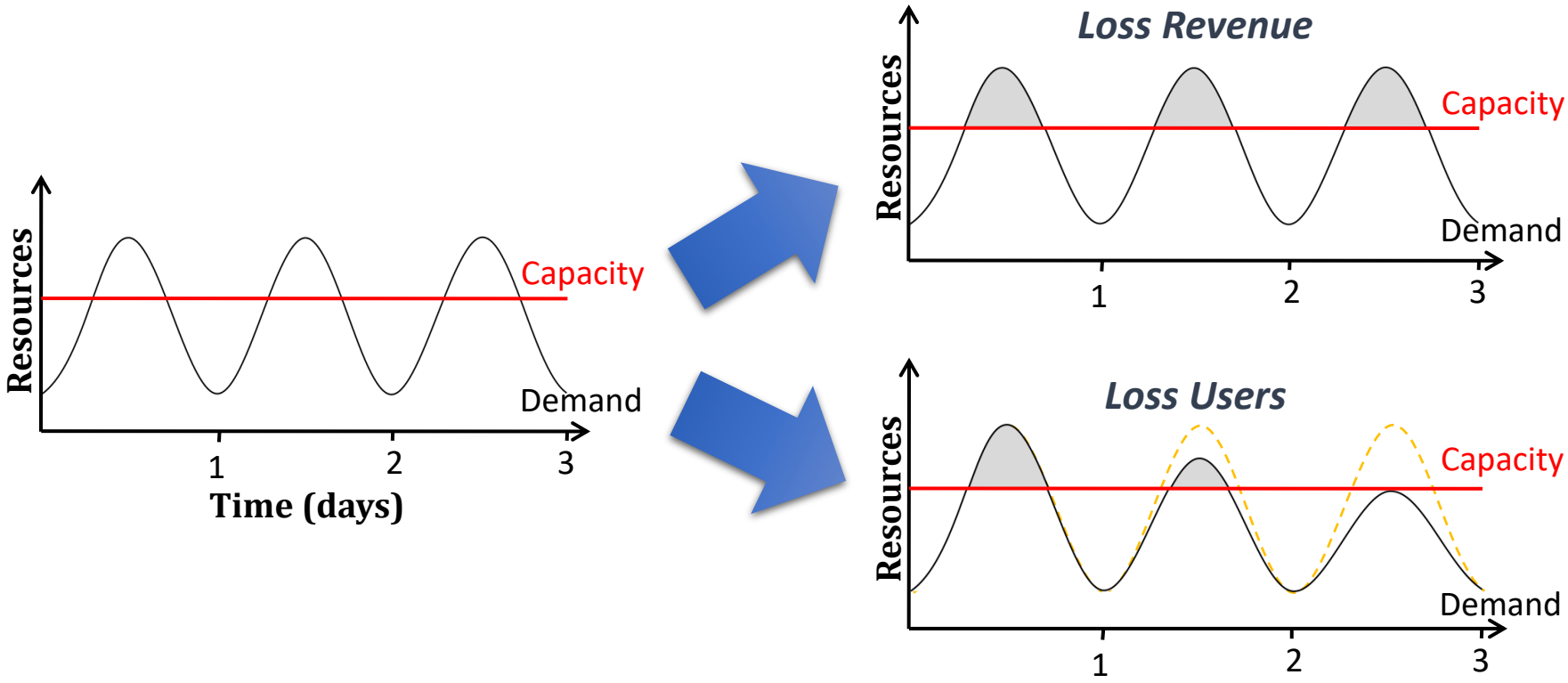
Dynamic Provisioning

- What is dynamic provisioning ?
  - Dynamic Provisioning is a simplified way to explain a complex networked server computing environment where server computing instances are provisioned or deployed from a administrative console or client application by the server administrator, network administrator, or any other enabled user.



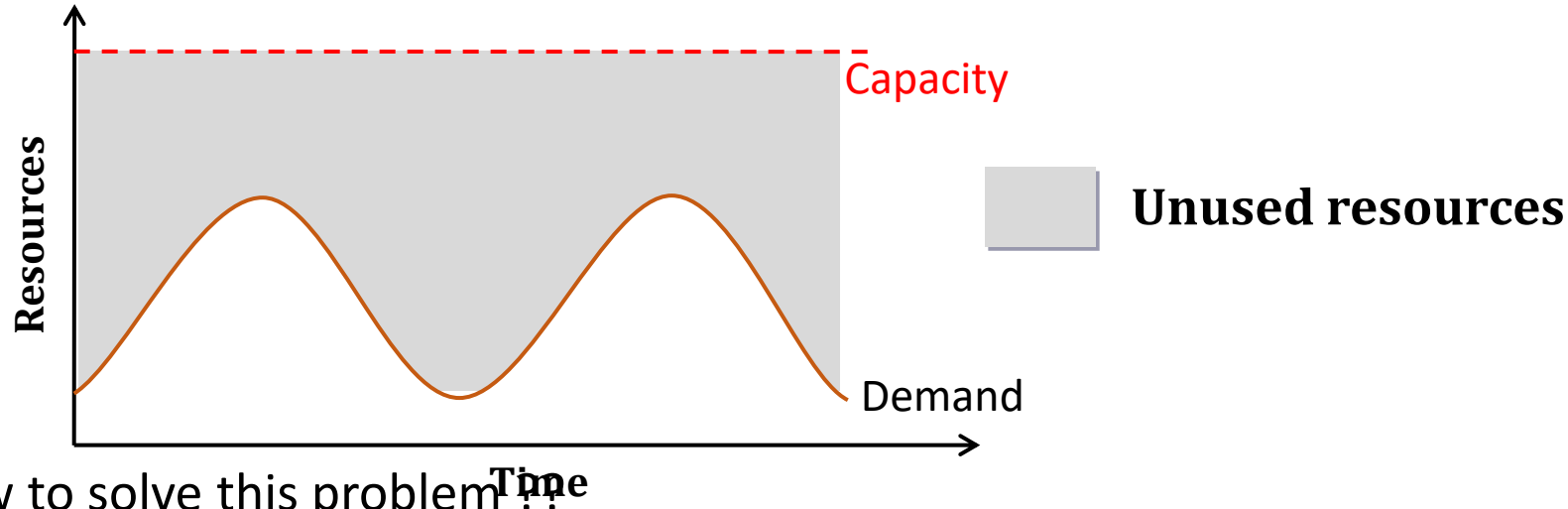
Dynamic Provisioning

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



Dynamic Provisioning

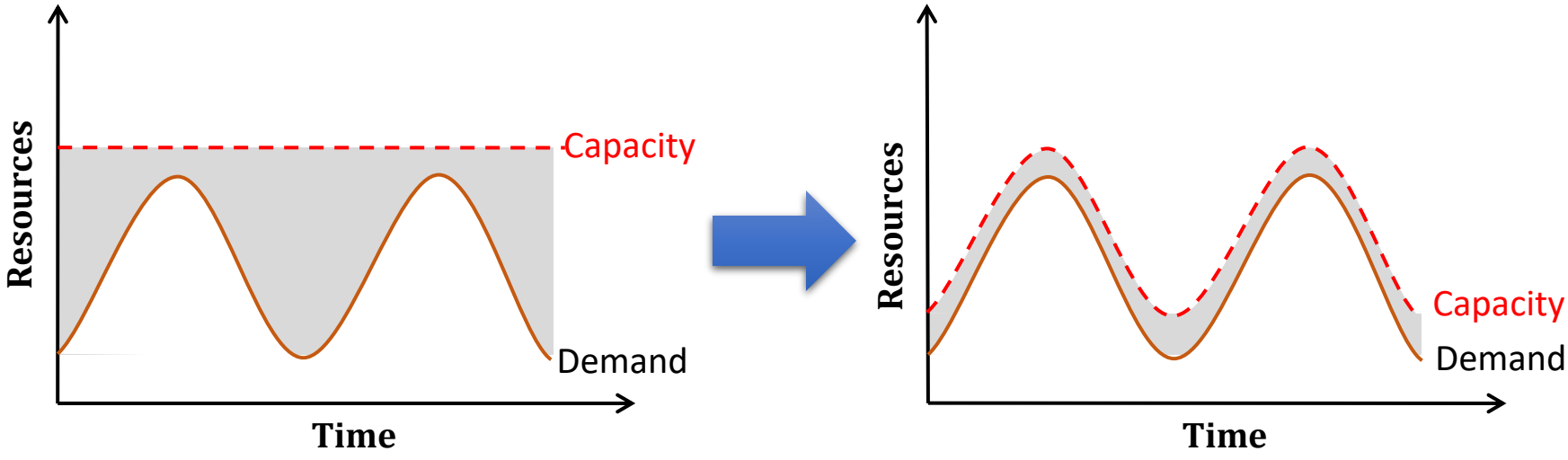
- Overestimate system utilization which result in low utilization



- How to solve this problem ?!
  - Dynamically provision resources

Dynamic Provisioning

- Cloud resources should be provisioned dynamically
  - Meet seasonal demand variations
  - Meet demand variations between different industries
  - Meet burst demand for some extraordinary events



Multi-tenant Design

- What is multi-tenant design ?
  - Multi-tenant refers to a principle in software architecture where a single instance of the software runs on a server, serving multiple client organizations.
  - With a multi-tenant architecture, a software application is designed to virtually partition its data and configuration thus each client organization works with a customized virtual application instance.

- Client oriented requirements :
  - Customization
    - Multi-tenant applications are typically required to provide a high degree of customization to support each target organization's needs.
  - Quality of service
    - Multi-tenant applications are expected to provide adequate levels of security and robustness.

Availability & Reliability

Data Never Loss

Machine Never Fail

Availability & Reliability

- What is availability ?
  - The degree to which a system, subsystem, or equipment is in a specified operable and committable state at the start of a mission, when the mission is called for at an unknown time.
  - Cloud system usually require high availability
    - Ex. “Five Nines” system would statistically provide 99.999% availability
- What is reliability ?
  - The ability of a system or component to perform its required functions under stated conditions for a specified period of time.
- But how to achieve these properties ?
  - Fault tolerance system
  - Require system resilience
  - Reliable system security

Fault Tolerance

- What is fault tolerant system ?
  - Fault-tolerance is the property that enables a system to continue operating properly in the event of the failure of some of its components.
  - If its operating quality decreases at all, the decrease is proportional to the severity of the failure, as compared to a naively-designed system in which even a small failure can cause total breakdown.

- Four basic characteristics :
  - No single point of failure
  - Fault detection and isolation to the failing component
  - Fault containment to prevent propagation of the failure
  - Availability of reversion modes