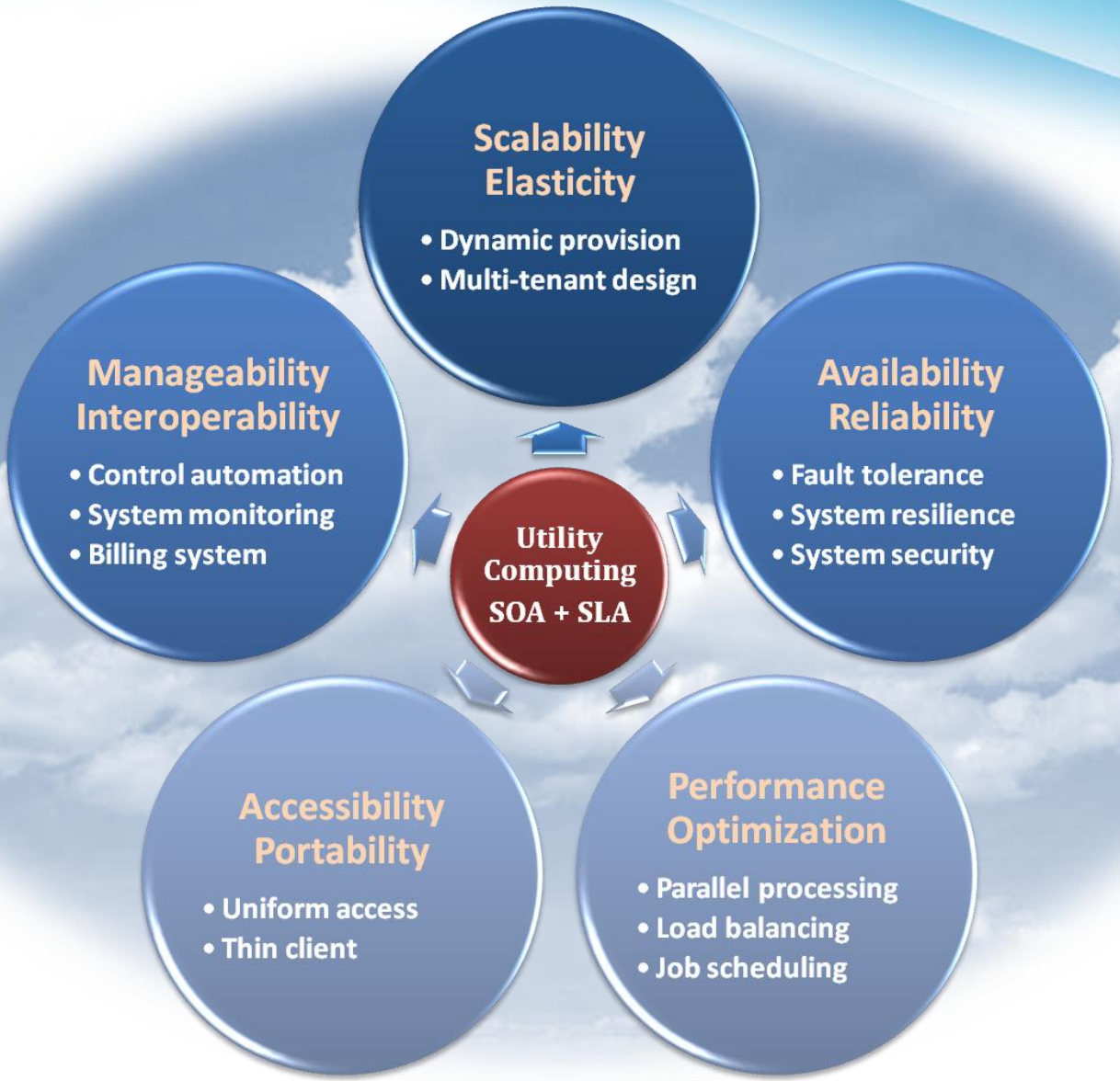
Properties and Characteristics



Manageability & Interoperability

- What is manageability ?
  - Enterprise-wide administration of cloud computing systems. Systems manageability is strongly influenced by network management initiatives in telecommunications.
- What is interoperability ?
  - Interoperability is a property of a product or system, whose interfaces are completely understood, to work with other products or systems, present or future, without any restricted access or implementation.
- But how to achieve these properties ?
  - System control automation
  - System state monitoring

Control Automation

- What is Autonomic Computing ?
  - Its ultimate aim is to develop computer systems capable of self-management, to overcome the rapidly growing complexity of computing systems management, and to reduce the barrier that complexity poses to further growth.

Control Automation

- Four functional areas :
  - Self-Configuration
    - Automatic configuration of components.
  - Self-Healing
    - Automatic discovery, and correction of faults.
  - Self-Optimization
    - Automatic monitoring and control of resources to ensure the optimal functioning with respect to the defined requirements.
  - Self-Protection
    - Proactive identification and protection from arbitrary attacks.

Manageability
Interoperability
- Control automation
- System monitoring
- Billing system

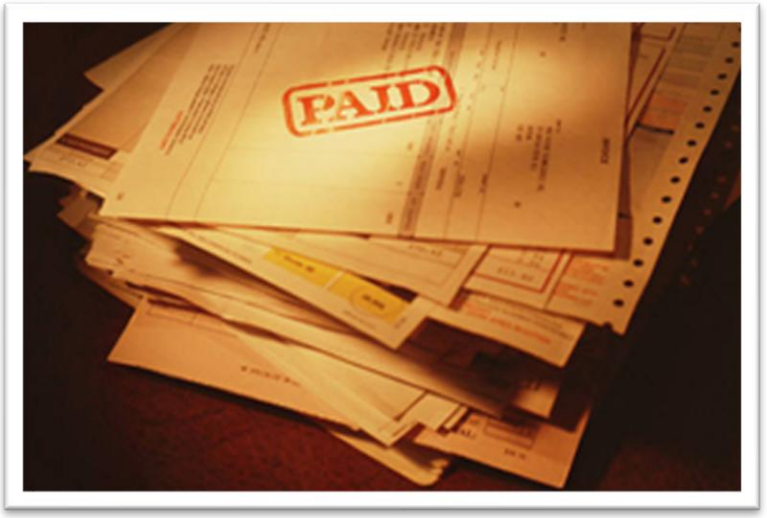
System Monitoring

- What is system monitor ?
  - A System Monitor in systems engineering is a process within a distributed system for collecting and storing state data.
- What should be monitored in the Cloud ?
  - Physical and virtual hardware state
  - Resource performance metrics
  - Network access patterns
  - System logs
  - … etc
- Anything more ?
  - Billing system

Manageability
Interoperability
- Control automation
- System monitoring
- Billing system

Billing System

- Billing System in Cloud
  - Users pay as many as they used.
  - Cloud provider must first determine the list of service usage price.
  - Cloud provider have to record the resource or service usage of each user, and then charge users by these records.
- How can cloud provider know users' usage ?
  - Get those information by means of monitoring system.
  - Automatically calculate the total amount of money which user should pay. And automatically request money from use's banking account.



Performance & Optimization

- Performance guarantees ??
  - As the great computing power in cloud, application performance should be guaranteed.
  - Cloud providers make use of powerful infrastructure or other underlining resources to build up a highly performed and highly optimized environment, and then deliver the complete services to cloud users.
- But how to achieve this property ?
  - Parallel computing
  - Load balancing
  - Job scheduling

Parallel Processing

- Parallel Processing
  - Parallel processing is a form of computation in which many calculations are carried out simultaneously, operating on the principle that large problems can often be divided into smaller ones, which are then solved concurrently.

- Parallelism in different levels :
  - Bit level parallelism
  - Instruction level parallelism
  - Data level parallelism
  - Task level parallelism

Performance Optimization
- Parallel processing
- Load balancing
- Job scheduling

Parallel Processing

- Hardware approaches
  - Multi-core computer
  - Symmetric multi-processor
  - General purpose graphic processing unit
  - Distributed computing
- Software approaches
  - Parallel programming language
  - Automatic parallelization

Performance Optimization
- Parallel processing
- Load balancing
- Job scheduling

Load Balancing

- What is load balancing ?
  - Load balancing is a technique to distribute workload evenly across two or more computers, network links, CPUs, hard drives, or other resources, in order to get optimal resource utilization, maximize throughput, minimize response time, and avoid overload.

- Why should be load balanced ?
  - Improve resource utilization
  - Improve system performance
  - Improve energy efficiency

Unbalanced ⟹

Job Scheduling

- What is job scheduler ?
  - A job scheduler is a software application that is in charge of unattended background executions, commonly known for historical reasons as batch processing.
- What should be scheduled in Cloud ?
  - Computation intensive tasks
  - Dynamic growing and shrinking tasks
  - Tasks with complex processing dependency
- How to approach ?
  - Use pre-defined workflow
  - System automatic configuration

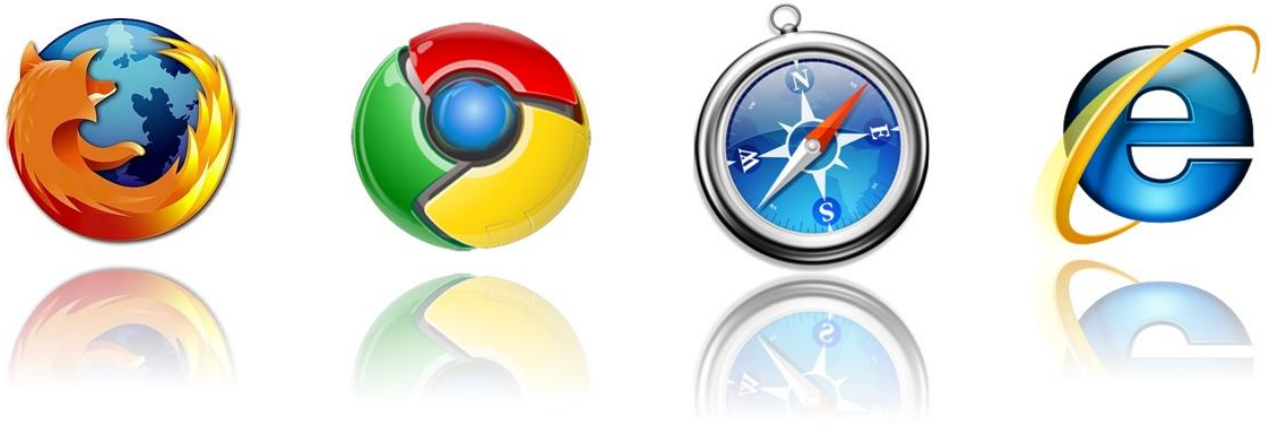
Accessibility & Portability

- What is accessibility ?
  - Accessibility is a general term used to describe the degree to which a product, device, service, or environment is accessible by as many people as possible.
- What is service portability ?
  - Service portability is the ability to access services using any devices, anywhere, continuously with mobility support and dynamic adaptation to resource variations.
- But how to achieve these properties ?
  - Uniform access
  - Thin client

Accessibility
Portability
• Uniform access
• Thin client

Uniform Access

- How do users access cloud services ?
  - Cloud provider should provide their cloud service by means of widespread accessing media. In other word, users from different operating systems or other accessing platforms should be able to directly be served.
  - Nowadays, web browser technique is one of the most widespread platform in almost any intelligent electronic devices. Cloud service take this into concern, and delivery their services with web-based interface through the Internet.



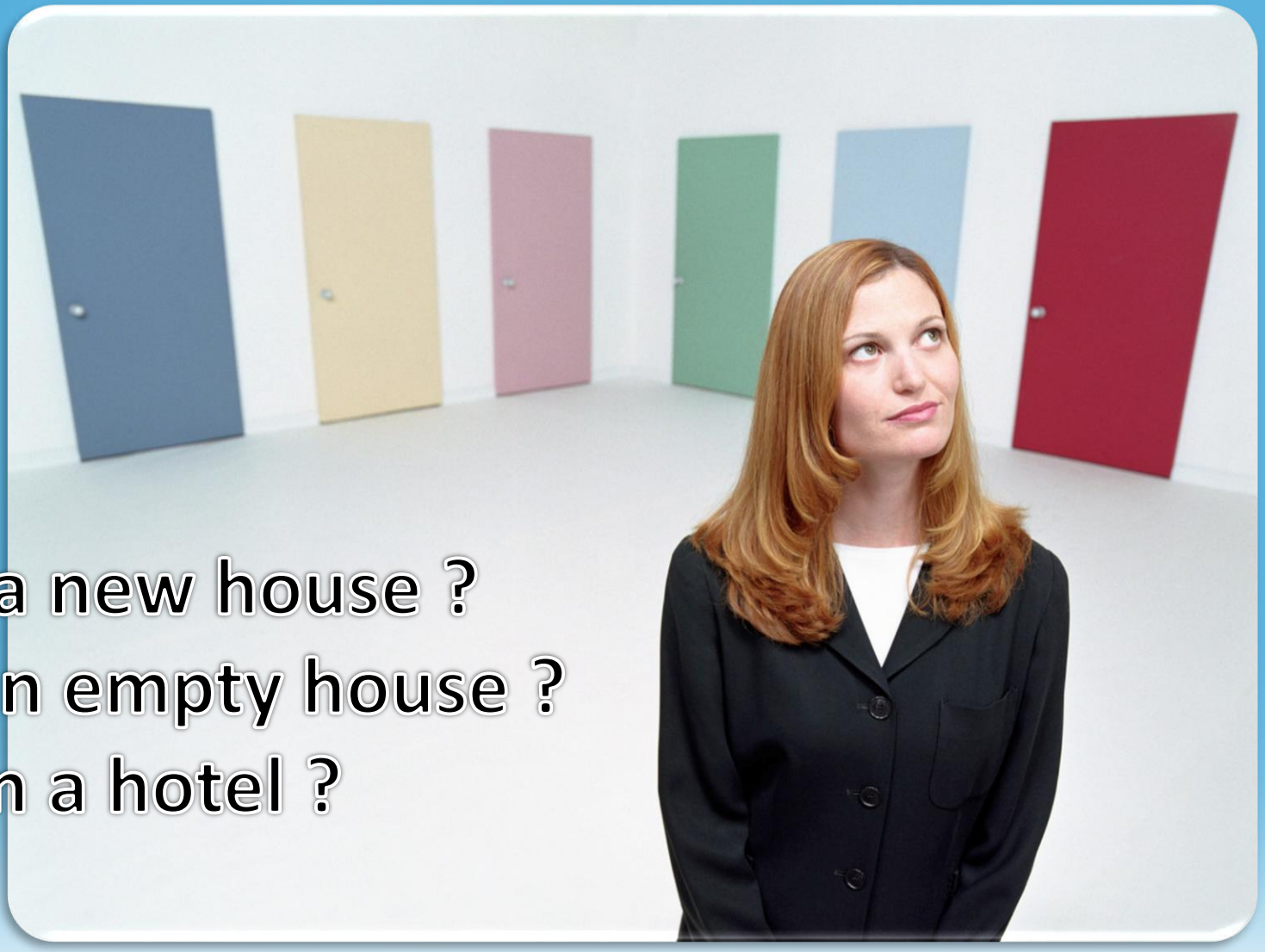
Thin Client

- What is thin client ?
  - Thin client is a computer or a computer program which depends heavily on some other computer to fulfill its traditional computational roles. This stands in contrast to the traditional fat client, a computer designed to take on these roles by itself.
- Characteristics :
  - Cheap client hardware
    - While the cloud providers handle several client sessions at once, the clients can be made out of much cheaper hardware.
  - Diversity of end devices
    - End user can access cloud service via plenty of various electronic devices, which include mobile phones and smart TV.
  - Client simplicity
    - Client local system do not need complete operational functionalities.

A Simple Analogy

Say, you just moved to a city and you are looking for a place to live.

What is your choice ?



Built a new house ?

Buy an empty house ?

Live in a hotel ?

Let’s built a new house !!

You can fully control everything your like your new house to have. But that is a hard work …

If you buy an empty house ?

You can customize some part of your house. But never change the original architecture.

How about live in a hotel ?

Live in a hotel will be a good idea if the only thing you care is enjoy your life!! There is nothing you can do with the house except living in it.

Let’s translate to Cloud Computing !!

Service Models Overview

- What if you want to have an IT department ?
  - Similar to ***build a new house*** in previous analogy
    - You can rent some virtualized infrastructure and build up your own IT system among those resources, which may be fully controlled.
    - Technical speaking, use the ***Infrastructure as a Service (IaaS)*** solution.
  - Similar to ***buy an empty house*** in previous analogy
    - You can directly develop your IT system through one cloud platform, and do not care about any lower level resource management.
    - Technical speaking, use the ***Platform as a Service (PaaS)*** solution.
  - Similar to ***live in a hotel*** in previous analogy
    - You can directly use some existed IT system solutions, which were provided by some cloud application service provider, without knowing any detail technique about how these service was achieved.
    - Technical speaking, use the ***Software as a Service (SaaS)*** solution.