Murang'a University of Technology

Innovation for Prosperity

DEPARTMENT OF INFORMATION TECHNOLOGY

SIT407: CLOUD COMPUTING

Credit hours: 3 hours

Course Notes

Topic: Migrating into a Cloud & Service Level Agreements

Lecturer's Name: J. Njuki

Presentation Outline

- Objectives
- Migrating into the cloud
- Migration Risks and Mitigation
- The challenges of SaaS paradigm
- Characteristics of integration solutions and products.
- Virtual machines provisioning and migration services analogy for virtual machine provisioning
- Live migration and high availability
- Service Level Agreement
- Cloud Computing Management and Application Trends
- Business Benefits of Cloud Computing

Objectives

1) Explain the seven-step model of migration into a cloud
2) Identify migration risks and mitigation
3) Describe the challenges of SaaS paradigm
4) Identify characteristics of integration solutions and products.
5) Explain virtual machines provisioning and migration services analogy for virtual machine provisioning
6) Describe live migration and high availability
7) Explain Service Level Agreement
8) Identify Cloud Computing Management and Application Trends
9) Explain Business Benefits of Cloud Computing

Migrating into the Cloud

The promise of cloud computing has raised the IT expectations of small and medium enterprises beyond measure. Large companies are deeply debating it. Cloud computing is a disruptive model of IT whose innovation is part technology and part business model in short a disruptive techno-commercial model of IT.

We propose the following definition of cloud computing: It is a techno-business disruptive model of using distributed large-scale data centers either private or public or hybrid offering customers a scalable virtualized infrastructure or an abstracted set of services qualified by service-level agreements (SLAs) and charged only by the abstracted IT resources consumed.

Migrating into the Cloud Cont'd

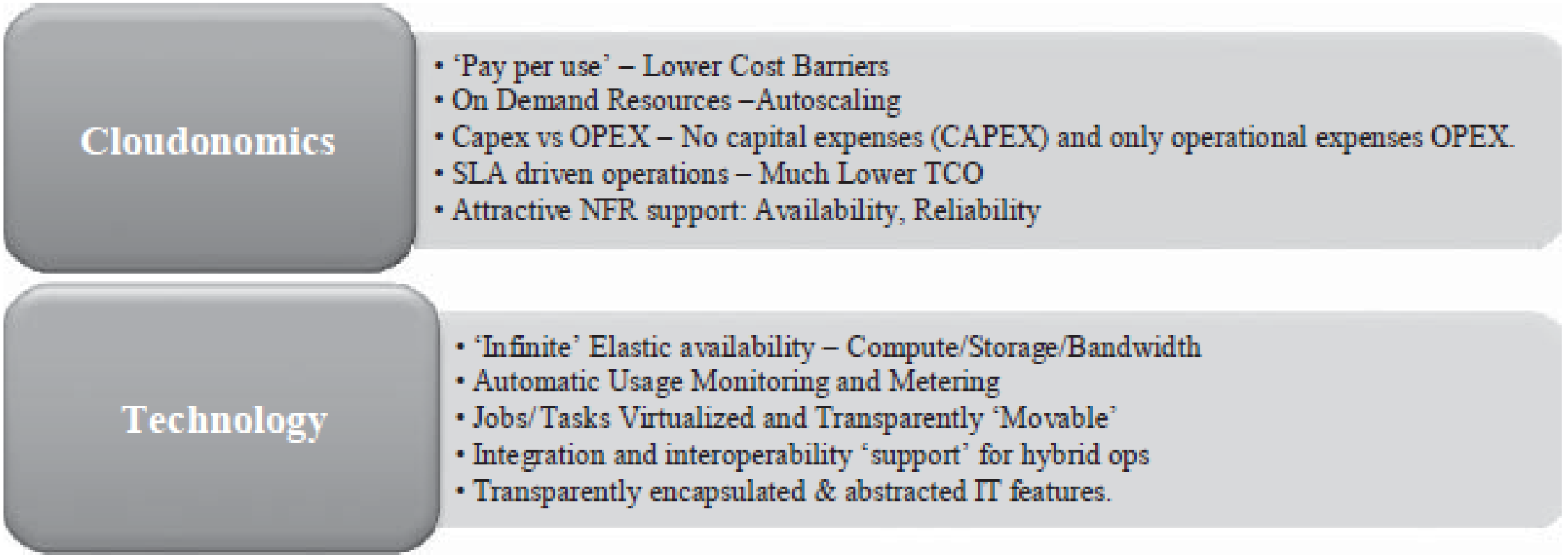


Figure 1. The promise of the cloud computing services

Migrating into the Cloud

In figure 1, the promise of the cloud both on the business front (the attractive cloudonomics) and the technology front widely aided the CxOs to spawn out several non-mission critical IT needs from the ambit of their captive traditional data centers to the appropriate cloud service.

Several small and medium business enterprises, however, leveraged the cloud much beyond the cautious user. Many startups opened their IT departments exclusively using cloud services very successfully and with high ROI. Having observed these successes, several large enterprises have started successfully running pilots for leveraging the cloud.

Migrating into the Cloud Cont'd

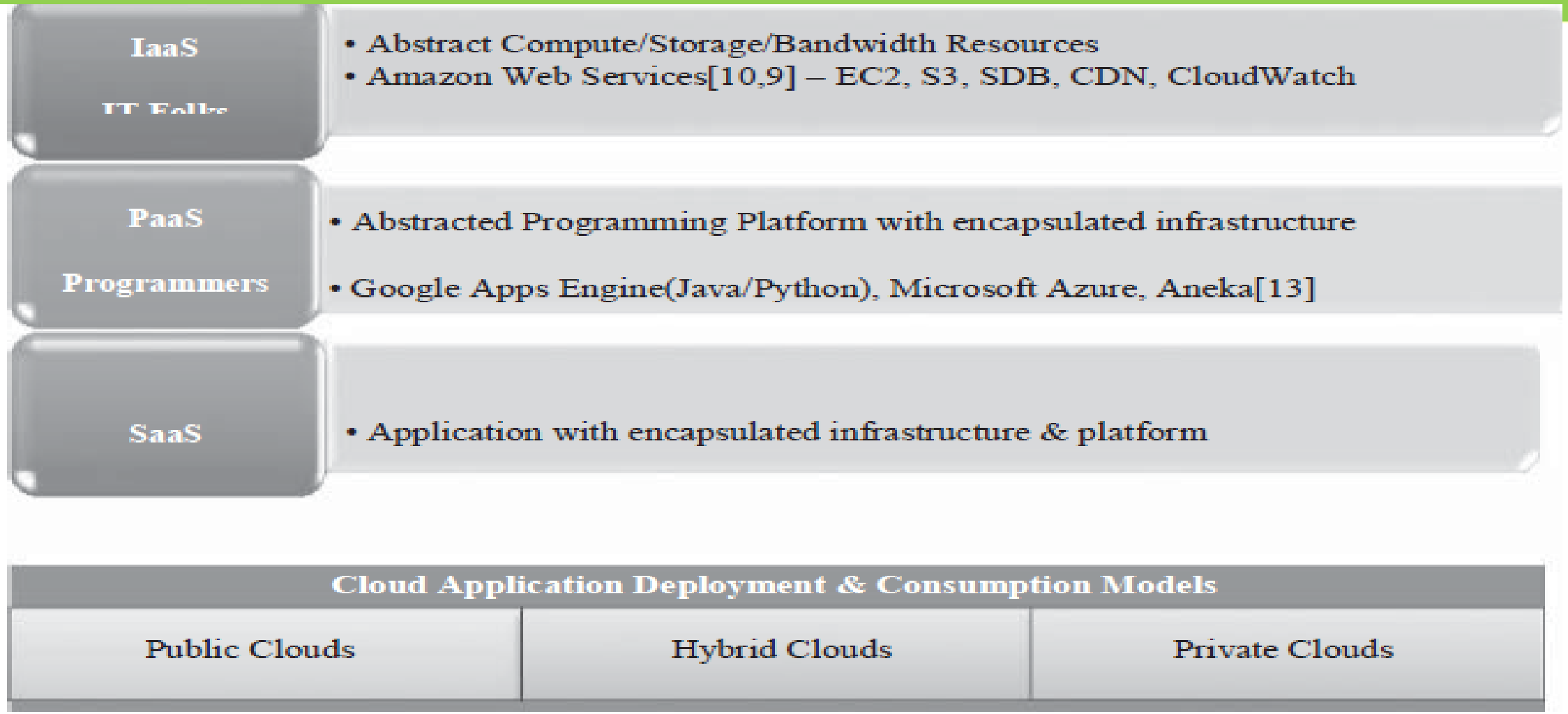


Figure 2: The cloud computing service offering and deployment models.

Migrating into the Cloud Cont'd

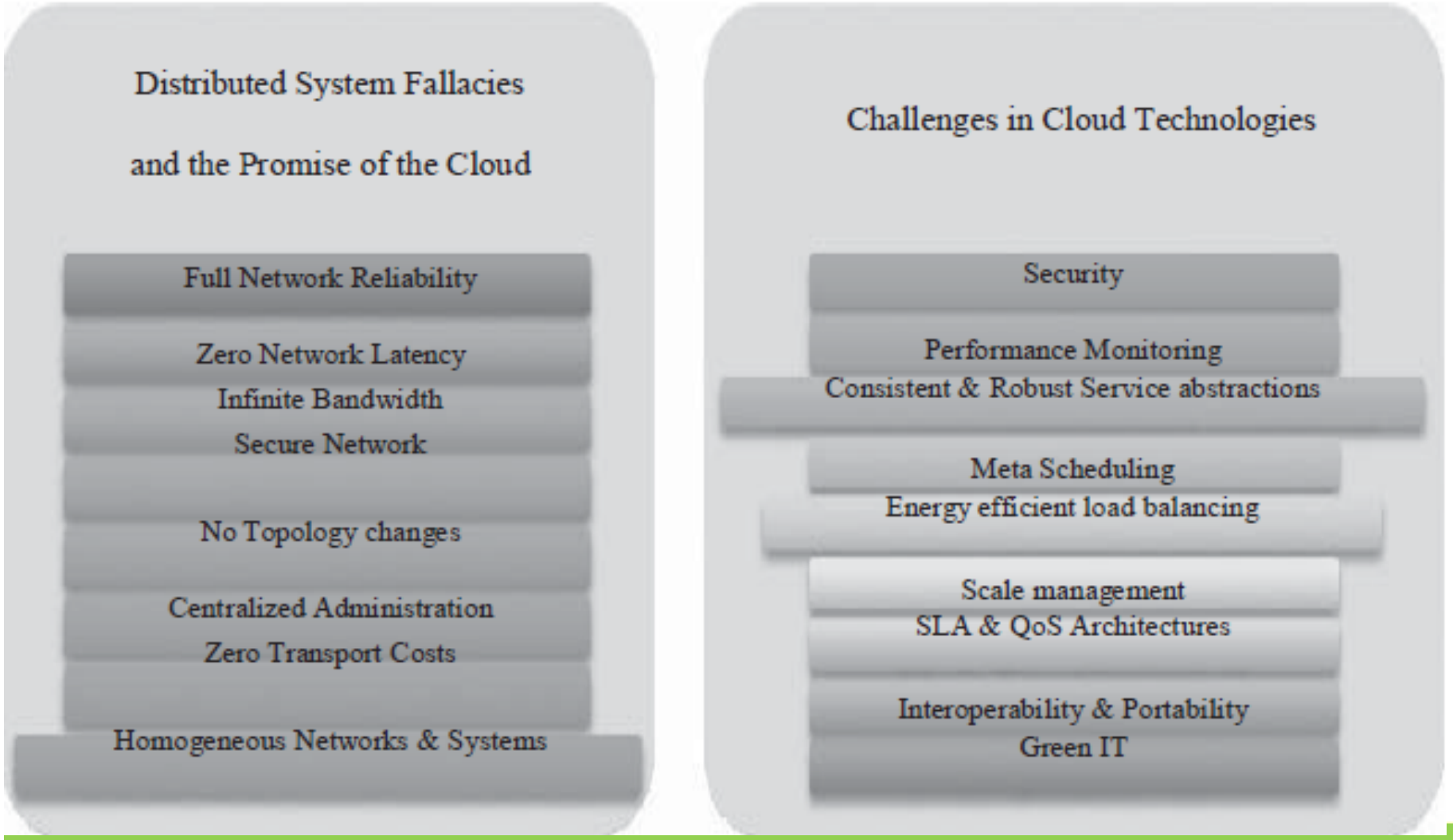


Figure 3: 'Under the hood' challenges of the cloud computing services implementations.

Migrating into the Cloud Cont'd

Broad approaches to migrating into the cloud

Cloud Economics deals with the economic rationale for leveraging the cloud and is central to the success of cloud-based enterprise usage.

Decision-makers, IT managers, and software architects are faced with several dilemmas when planning for new Enterprise IT initiatives

The seven-step model of migration into a cloud

Typically migration initiatives into the cloud are implemented in phases or in stages. A structured and process-oriented approach to migration into a cloud has several advantages of capturing within itself the best practices of many migration projects

Migrating into the Cloud Cont'd

1. Conduct Cloud Migration Assessments
2. Isolate the Dependencies
3. Map the Messaging & Environment
4. Re-architect & Implement the lost Functionalities
5. Leverage Cloud Functionalities & Features
6. Test the Migration
7. Iterate and Optimize

Migrating into the Cloud Cont'd

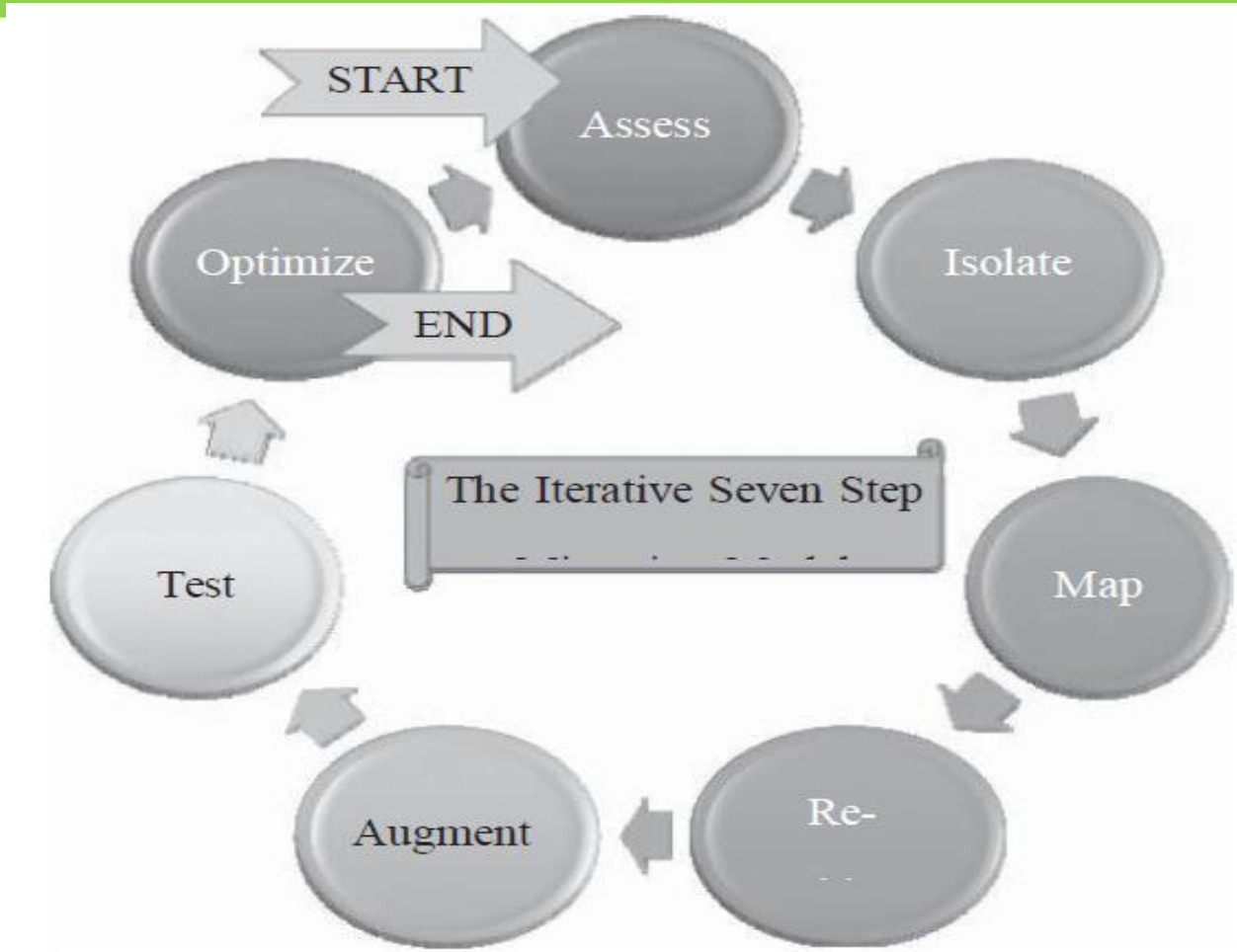


Figure 4: The iterative Seven-step Model of Migration into the Cloud

Migrating into the Cloud Cont'd

1) Assessment:-

The first step of the iterative process of the seven-step model of migration is basically at the assessment level. These assessments are about the cost of migration as well as about the ROI that can be achieved in the case of production version.

Isolating:-

The next process step is in isolating all systemic and environmental dependencies of the enterprise application components within the captive data center. This, in turn, yields a picture of the level of complexity of the migration.

Migrating into the Cloud Cont'd

2) Mapping:-

After isolation is complete, one then goes about generating the mapping constructs between what shall possibly remain in the local captive data center and what goes onto the cloud.

The Seven-Step Model Of Migration Into A Cloud

3) Re-Architect:-

Perhaps a substantial part of the enterprise application needs to be re-architected, redesigned, and re-implemented on the cloud. This gets in just about the functionality of the original enterprise application. Due to this migration, it is possible perhaps that some functionality is lost.

Migrating into the Cloud Cont'd

5) Augment:-

In the next process step we leverage the internal features of the cloud computing service to augment our enterprise application in its own small ways.

6) Validate:-

Having done the augmentation, we validate and test the new form of the enterprise application with an extensive test suite that comprises testing the components of the enterprise application on the cloud as well. These test results could be positive or mixed.

7) Optimization:-

In the latter case, we iterate and optimize as appropriate. After several such optimizing iterations, the migration is deemed successful

Migration Risks and Mitigation

The biggest challenge to any cloud migration project is how effectively the migration risks are identified and mitigated. In the Seven-Step Model of Migration into the Cloud, the process step of testing and validating includes efforts to identify the key migration risks.

In the optimization step, we address various approaches to mitigate the identified migration risks.

Migration risks for migrating into the cloud fall under two broad categories: the general migration risks and the security-related migration risks.

Migration Risks and Mitigation Cont'd

In the former we address several issues including

- performance monitoring and tuning essentially identifying all possible production level deviants;
- the business continuity and disaster recovery in the world of cloud computing service;
- the compliance with standards and governance issues; the IP and licensing issues;
- the quality of service (QoS) parameters as well as the corresponding SLAs committed to;
- the ownership, transfer, and storage of data in the application; the portability and interoperability issues which could help mitigate potential vendor lock-ins;

Migration Risks and Mitigation Cont'd

- the issues that result in trivializing and non comprehending the complexities of migration that results in migration failure and loss of senior management's business confidence in these efforts.

As with any new technology, SaaS and cloud concepts too suffer a number of limitations.

These technologies are being diligently examined for specific situations and scenarios.

The prickling and tricky issues in different layers and levels are being looked into. The overall views are listed out

Migration Risks and Mitigation Cont'd

Loss or lack of the following features deters the massive adoption of clouds

1. Controllability
2. Visibility & flexibility
3. Security and Privacy
4. High Performance and Availability
5. Integration and Composition
6. Standards

A number of approaches are being investigated for resolving the identified issues and flaws. Private cloud, hybrid and the latest community cloud are being prescribed as the solution for most of these inefficiencies and deficiencies.

Migration Risks and Mitigation Cont'd

As rightly pointed out by someone in his weblogs, still there are miles to go. There are several companies focusing on this issue Integration Conundrum. While SaaS applications offer outstanding value in terms of features and functionalities relative to cost, they have introduced several challenges specific to integration. The first issue is that the majority of SaaS applications are point solutions and service one line of business.

APIs are Insufficient: Many SaaS providers have responded to the integration challenge by developing application programming interfaces (APIs). Unfortunately, accessing and managing data via an API requires a significant amount of coding as well as maintenance due to frequent API modifications and updates.

Migration Risks and Mitigation Cont'd

Data Transmission Security: SaaS providers go to great length to ensure that customer data is secure within the hosted environment. However, the need to transfer data from on-premise systems or applications behind the firewall with SaaS applications hosted outside of the client's data center poses new challenges that need to be addressed by the integration solution of choice.

The Impacts of Cloud:. On the infrastructural front, in the recent past, the clouds have arrived onto the scene powerfully and have extended the horizon and the boundary of business applications, events and data.

Migration Risks and Mitigation Cont'd

That is, business applications, development platforms etc. are getting moved to elastic, online and on-demand cloud infrastructures. Precisely speaking, increasingly for business, technical, financial and green reasons, applications and services are being readied and relocated to highly scalable and available clouds.

The Integration Methodologies

Excluding the custom integration through hand-coding, there are three types for cloud integration

1. Traditional Enterprise Integration Tools can be empowered with special connectors to access Cloud-located Applications.

This is the most likely approach for IT organizations, which have already invested a lot in integration suite for their application integration needs.

Migration Risks and Mitigation Cont'd

2. Traditional Enterprise Integration Tools are hosted in the Cloud

This approach is similar to the first option except that the integration software suite is now hosted in any third-party cloud infrastructures so that the enterprise does not worry about procuring and managing the hardware or installing the integration software. This is a good fit for IT organizations that outsource the integration projects to IT service organizations and systems integrators, who have the skills and resources to create and deliver integrated systems.

3. Integration-as-a-Service (IaaS) or On-Demand Integration Offerings

Migration Risks and Mitigation Cont'd

These are SaaS applications that are designed to deliver the integration service securely over the Internet and are able to integrate cloud applications with the on-premise systems, cloud-to-cloud applications. SaaS administrator or business analyst as the primary resource for managing and maintaining their integration work. A good example is Informatica On-Demand Integration Services.

In the integration requirements can be realized using any one of the following methods and middleware products.

1. Hosted and extended ESB (Internet service bus / cloud integration bus)

Migration Risks and Mitigation Cont'd

2. Online Message Queues, Brokers and Hubs
3. Wizard and configuration-based integration platforms (Niche integration solutions)
4. Integration Service Portfolio Approach
5. Appliance-based Integration (Standalone or Hosted)

The key attributes of integration platforms and backbones gleaned and gained from integration projects experience are connectivity, semantic mediation, Data mediation, integrity, security, governance etc

- Connectivity refers to the ability of the integration engine to engage with both the source and target systems using available native interfaces.

Migration Risks and Mitigation Cont'd

This means leveraging the interface that each provides, which could vary from standards-based interfaces, such as Web services, to older and proprietary interfaces. Systems that are getting connected are very much responsible for the externalization of the correct information and the internalization of information once processed by the integration engine.

- Semantic Mediation refers to the ability to account for the differences between application semantics between two or more systems.

Semantics means how information gets understood, interpreted and represented within information systems. When two different and distributed systems are linked, the differences between their own yet distinct semantics have to be covered.

Migration Risks and Mitigation Cont'd

- Data Mediation converts data from a source data format into destination data format. Coupled with semantic mediation, data mediation or data transformation is the process of converting data from one native format on the source system, to another data format for the target system.
- Data Migration is the process of transferring data between storage types, formats, or systems. Data migration means that the data in the old system is mapped to the new systems, typically leveraging data extraction and data loading technologies.

Virtual machines provisioning and migration services analogy for virtual machine provisioning:

Historically, when there is a need to install a new server for a certain workload to provide a particular service for a client,

Migration Risks and Mitigation Cont'd

lots of effort was exerted by the IT administrator, and much time was spent to install and provision a new server.

1) Check the inventory for a new machine,
2) get one,
3) format, install OS required,
4) and install services; a server is needed along with lots of security batches and appliances.

Now, with the emergence of virtualization technology and the cloud computing IaaS model:

- It is just a matter of minutes to achieve the same task. All you need is to provision a virtual server through a self-service interface with small steps to get what you desire with the required specifications.

Migration Risks and Mitigation Cont'd

1) provisioning this machine in a public cloud like Amazon Elastic Compute Cloud (EC2), or
2) using a virtualization management software package or a private cloud management solution installed at your data center in order to provision the virtual machine inside the organization and within the private cloud setup. Analogy for Migration Services:

- Previously, whenever there was a need for performing a server's upgrade or performing maintenance tasks, you would exert a lot of time and effort, because it is an expensive operation to maintain or upgrade a main server that has lots of applications and users.

Migration Risks and Mitigation Cont'd

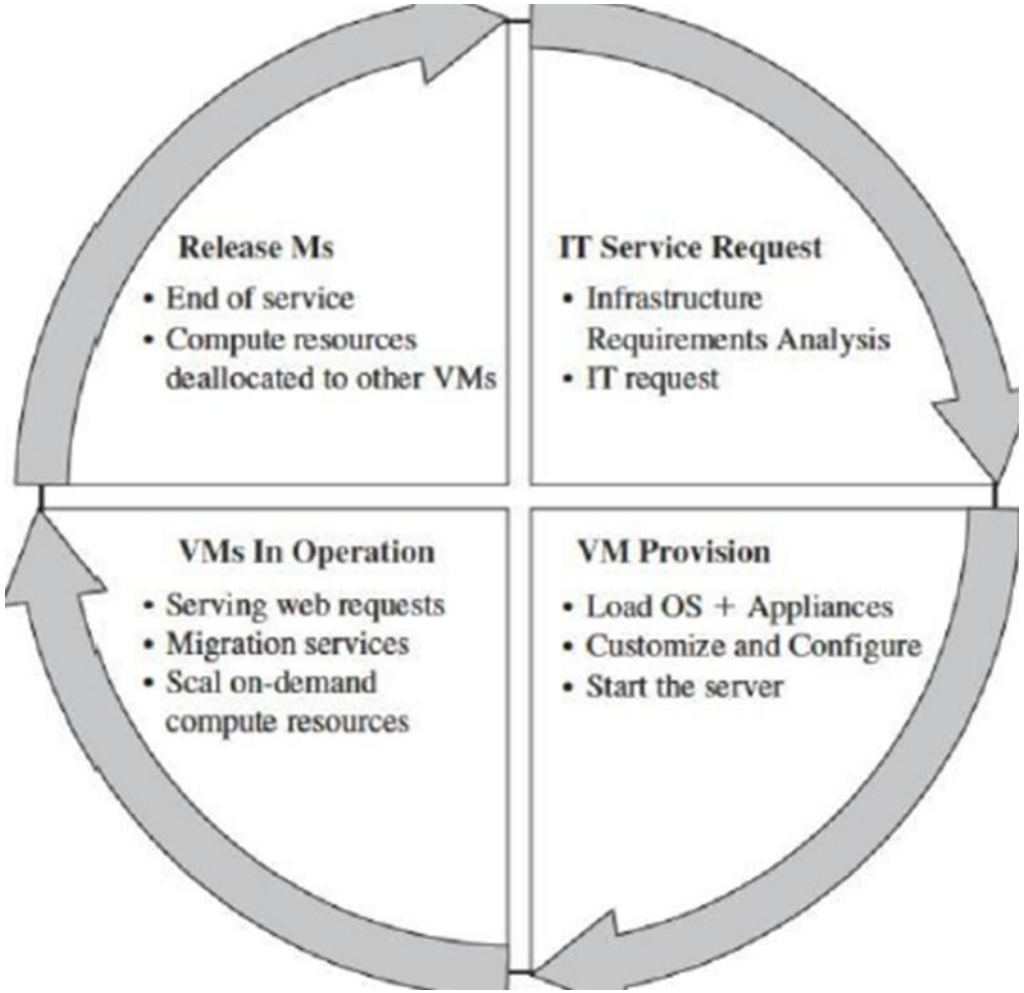
- Now, with the advance of the revolutionized virtualization technology and migration services associated with hypervisors' capabilities, these tasks (maintenance, upgrades, patches, etc.) are very easy and need no time to accomplish.
- Provisioning a new virtual machine is a matter of minutes, saving lots of time and effort, Migrations of a virtual machine is a matter of millisecondsVirtual Machine Provisioning and Manageability

Virtual machine life cycle

The cycle starts by a request delivered to the IT department, stating the requirement for creating a new server for a particular service.

Migration Risks and Mitigation Cont'd

- This request is being processed by the IT administration to start seeing the servers' resource pool, matching these resources with requirements
- Starting the provision of the needed virtual machine.
- Once it provisioned and started, it is ready to provide the required service according to an SLA.
- Virtual is being released; and free resources.



VM PROVISIONING PROCESS

The common and normal steps of provisioning a virtual server are as follows:

- Firstly, you need to select a server from a pool of available servers (physical servers with enough capacity) along with the appropriate OS template you need to provision the virtual machine.
- Secondly, you need to load the appropriate software (operating System you selected in the previous step, device drivers, middleware, and the needed applications for the service required).
- Thirdly, you need to customize and configure the machine (e.g., IP address, Gateway) to configure an associated network and storage resources.
- Finally, the virtual server is ready to start with its newly loaded software.

VM Provisioning Process Cont'd

To summarize, server provisioning is defining server's configuration based on the organization requirements, a hardware, and software component (processor, RAM, storage, networking, operating system, applications, etc.).

- Normally, virtual machines can be provisioned by manually installing an operating system, by using a preconfigured VM template, by cloning an existing VM, or by importing a physical server or a virtual server from another hosting platform. Physical servers can also be virtualized and provisioned using P2V (Physical to Virtual) tools and techniques (e.g., virt- p2v).

VM Provisioning Process Cont'd

- After creating a virtual machine by virtualizing a physical server, or by building a new virtual server in the virtual environment, a template can be created out of it.
- Most virtualization management vendors (VMware, XenServer, etc.) provide the data center's administration with the ability to do such tasks in an easy way.

Live Migration and High Availability

Live migration (which is also called hot or real-time migration) can be defined as the movement of a virtual machine from one physical host to another while being powered on.

VIRTUAL MACHINE PROVISIONING PROCESS

DATA LOCK-IN AND STANDARDIZATION

The Cloud Computing Interoperability Forum (CCIF) was formed by organizations such as Intel, Sun, and Cisco in order to enable a global cloud computing ecosystem whereby organizations are able to seamlessly work together for the purposes for wider industry adoption of cloud computing technology.

The development of the Unified Cloud Interface (UCI) by CCIF aims at creating a standard programmatic point of access to an entire cloud infrastructure in the hardware virtualization sphere, the Open Virtual Format (OVF) aims at facilitating packing and distribution of software to be run on VMs so that virtual appliances can be made portable

DATA LOCK-IN AND STANDARDIZATION CONT'D

- Cloud lock-in (also known as vendor lock-in or data lock-in) occurs when transitioning data, products, or services to another vendor's platform is difficult and costly, making customers more dependent (locked-in) on a single cloud storage solution.
- The vendor lock-in problem in cloud computing is the situation where customers are dependent (i.e. locked-in) on a single cloud provider technology implementation and cannot easily move in the future to a different vendor without substantial costs, legal constraints, or technical incompatibilities
- A real-world example of vendor lock-in is the way Apple locked consumers into using iTunes in the early days of the service, because music purchased via iTunes could only be played within the iTunes application or on an iPod.

Monitoring and Management

Monitoring and management: an architecture for federated cloud computing the basic principles of cloud computing

In this section we unravel a set of principles that enable Internet scale cloud computing services.

We seek to highlight the fundamental requirement from the providers of cloud computing to allow virtual applications to freely migrate, grow, and shrink

Monitoring and Management Cont'd

Federation

- All cloud computing providers, regardless of how big they are, have a finite capacity. To grow beyond this capacity, cloud computing providers should be able to form federations of providers such that they can collaborate and share their resources.
- The need for federation-capable cloud computing offerings is also derived from the industry trend of adopting the cloud computing paradigm internally within companies to create private clouds and then being able to extend these clouds with resources leased on-demand from public clouds.

Monitoring and Management Cont'd

Independence

- Just as in other utilities, where we get service without knowing the internals of the utility provider and with standard equipment not specific to any provider (e.g., telephones), for cloud computing services to really fulfill the computing as a utility vision, we need to offer cloud computing users full independence.
- Users should be able to use the services of the cloud without relying on any provider- specific tool, and cloud computing providers should be able to manage their infrastructure without exposing internal details to their customers or partners.

Monitoring and Management Cont'd

provider- specific tool, and cloud computing providers should be able to manage their infrastructure without exposing internal details to their customers or partners. As a consequence of the independence principle, all cloud services need to be encapsulated and generalized such that users will be able to acquire equivalent virtual resources at different providers

Isolation

- Cloud computing services are, by definition, hosted by a provider that will simultaneously host applications from many different users.

Monitoring and Management Cont'd

For these users to move their computing into the cloud, they need warranties from the cloud computing provider that their stuff is completely isolated from others.

Users must be ensured that their resources cannot be accessed by others sharing the same cloud and that adequate performance isolation is in place to ensure that no other user may possess the power to directly effect the service granted to their application.

Monitoring and Management Cont'd

Elasticity

- One of the main advantages of cloud computing is the capability to provide, or release, resources on-demand.
- These elasticity capabilities should be enacted automatically by cloud computing providers to meet demand variations, just as electrical companies are able (under normal operational circumstances) to automatically deal with variances in electricity consumption levels.
- Clearly the behavior and limits of automatic growth and shrinking should be driven by contracts and rules agreed on between cloud computing providers and consumers.

Monitoring and Management Cont'd

Trust

- Probably the most critical issue to address before cloud computes that of establishing trust.
- Mechanisms to build and maintain trust between cloud computing consumers and cloud computing providers, as well as between cloud computing providers among themselves, are essential for the success of any cloud computing offeringing can become the preferred computing paradigm

Service Level Agreement

In an IaaS model it is expected from the service provider that it sizes capacity demands for its service. If resource demands are provided correctly and are indeed satisfied upon request , then desired user experience of the service will be guar anteed.

A risk mitigation mechanism to protect user experience in the IaaS model is offered by infrastructure SLAs (i.e., the SLAs formalizing capacity availability) signed between service provider and IaaS provider.

There are three main approaches

- No SLAs . This approach is based on two premises :

a) Cloud always has spare capacity to provide on demand, and

Service Level Agreement Cont'd

b) services are not QoS sensitive and can withstand moderate performance degradation. This methodology is best suited for the best effort workloads

- Probabilistic SLAs. These SLAs allow us to trade capacity availability for cost of consumption. Probabilistic SLAs specify clauses that determine availability percentile for contracted resources computed over the SLA evaluation period. The lower the availability percentile, the cheaper the cost of resource consumption
- Deterministic SLAs. These are, in fact, probabilistic SLAs where resource availability percentile is 100%. These SLAs are most stringent and difficult to guarantee .

Service Level Agreement Cont'd

- From the provider's point of view, they do not admit capacity multiplexing. Therefore this is the most costly option for service providers , which may be applied for critical services Elasticity rule s are scaling and de- scaling policies that guide transition of the service from one configuration to another to match change s in the environment . The main motivation for defining these policies stems from the pay- as you go billing model of IaaS clouds. The service owner is interested in paying only for what is really required to satisfy workload demands minimizing the overprovisioning overhead

Service Level Agreement Cont'd

There are three types of elasticity rules:

- Time-driven: These rules change the virtual resources array in response to a timer event . These rules are useful for predictable Workloads, for example, for services with well-known business cycles.
- OS Level Metrics-Driven: These rules react on predicates defined in terms of the OS parameters observable in the black box mode. These auto-scaling policies are useful for transparently scaling and de scaling services..
- Application Metrics –Driven: This is a unique RESERVOIR offering that allows an application to supply application- specific policies that will be transparently executed by IaaS middleware in reacting on the monitoring information supplied by the service- specific monitoring probe s running inside VMs.

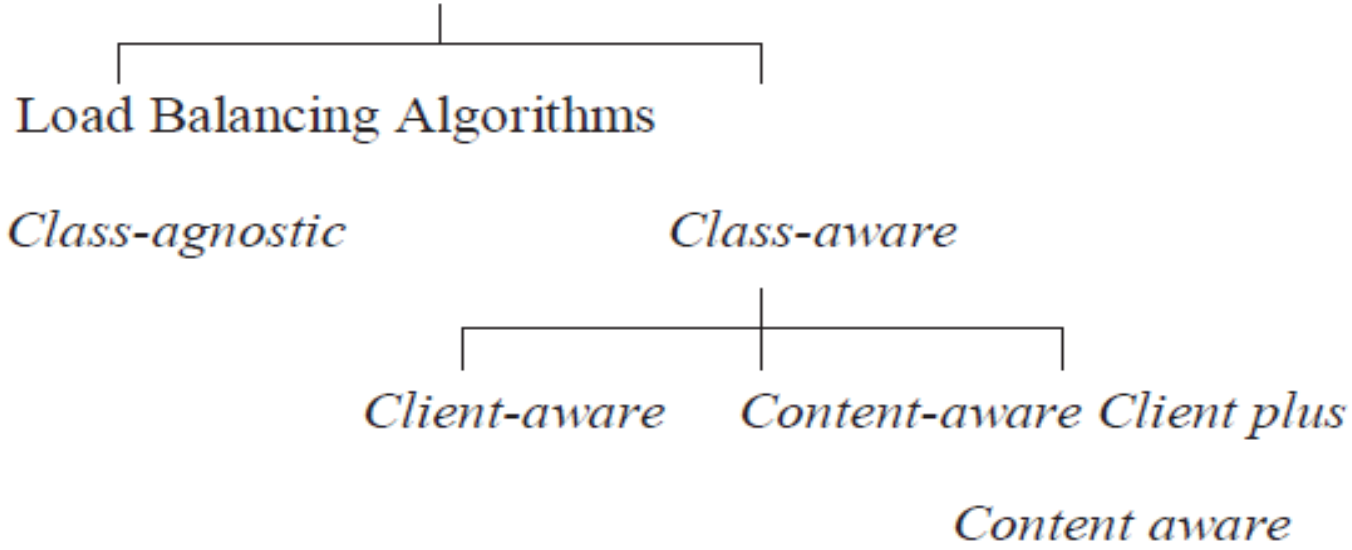
Service Level Agreement Cont'd

SLA Management In Cloud Computing: a service provider's perspective traditional approaches to SLO management traditionally, load balancing techniques and admission control mechanisms have been used to provide guaranteed quality of service (QoS) for hosted web applications.
These mechanisms can be viewed as the first attempt towards managing the SLOs. In the following subsections we discuss the existing approaches for load balancing and admission control for ensuring QoS.

Service Level Agreement Cont'd

Load Balancing

The objective of a load balancing is to distribute the incoming requests onto a set of physical machines, each hosting a replica of an application.



General taxonomy of load-balancing algorithms.

Service Level Agreement Cont'd

The load balancing algorithm executes on a physical machine that interfaces with the clients.

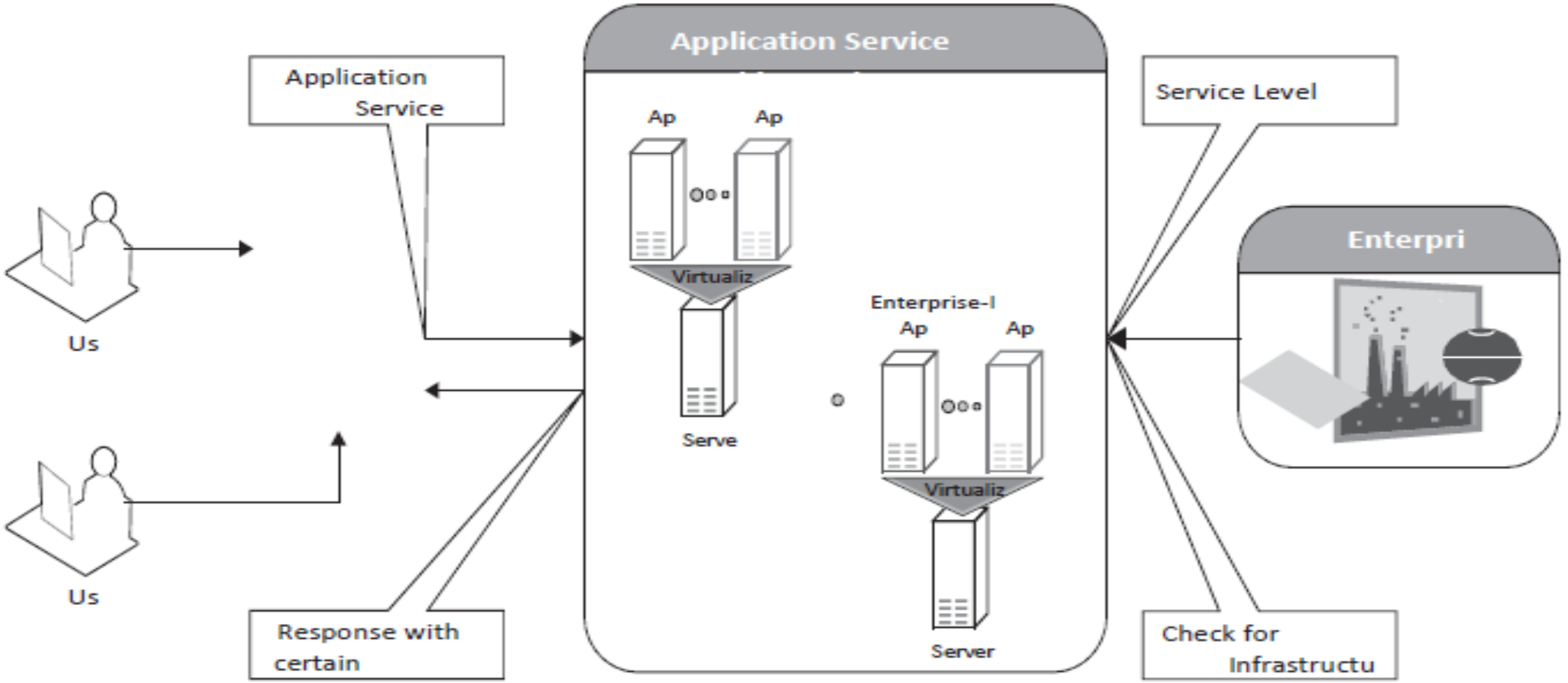
This physical machine, also called the front-end node, receives the incoming requests and distributes these requests to different physical machines for further execution.

This set of physical machines is responsible for serving the incoming requests and are known as the back-end nodes

Typically, the algorithm executing on the front-end node is agnostic to the nature of the request.

This means that the front-end node is neither aware of the type of client from which the request originates nor aware of the category

Service Level Agreement Cont'd



Shared hosting of applications on virtualized servers within ASP's data centers.

Service Level Agreement Cont'd

e.g., browsing, selling, payment, etc. to which the request belongs to. This category of load balancing algorithms is known as class-agnostic.

There is a second category of load balancing algorithms that is known as class-aware.

With class-aware load balancing and requests distribution, the front-end node must additionally inspect the type of client making the request and/or the type of service requested before deciding which back-end node should service the request.

Inspecting a request to find out the class or category of a request is difficult because the client must first establish a connection with a node (front-end node) that is not responsible for servicing the request.

Service Level Agreement Cont'd

Admission Control

Admission control algorithms play an important role in deciding the set of requests that should be admitted into the application server when the server experiences very heavy loads during overload situations, since the response time for all the requests would invariably degrade if all the arriving requests are admitted into the server, it would be preferable to be selective in identifying a subset of requests that should be admitted into the system so that the overall pay-off is high.

The objective of admission control mechanisms, therefore, is to police the incoming requests and identify a subset of incoming requests that can be admitted into the system when the system faces overload situations.

Service Level Agreement Cont'd

TYPES OF SLA

- Service-level agreement provides a framework within which both seller and buyer of a service can pursue a profitable service business relationship.
- It outlines the broad understanding between the service provider and the service consumer for conducting business and forms the basis for maintaining a mutually beneficial relationship.
- From a legal perspective, the necessary terms and conditions that bind the service provider to provide services continually to the service consumer are formally defined in SLA.

Service Level Agreement Cont'd

- SLA can be modeled using web service-level agreement (WSLA) language specification.
- Although WSLA is intended for web-service-based applications, it is equally applicable for hosting of applications.

Service-level parameter, metric, function, measurement directive, service-level objective, and penalty are some of the important components of WSLA and are described as follows

Key Components of a Service-Level Agreement

Service-Level Parameter - Describes an observable property of a service whose value is measurable.

Service Level Agreement Cont'd

Metrics - These are definitions of values of service properties that are measured from a service-providing system or computed from other metrics and constants. Metrics are the key instrument to describe exactly what SLA parameters mean by specifying how to measure or compute the parameter values.

Function - A function specifies how to compute a metric's value from the values of other metrics and constants. Functions are central to describing exactly how SLA parameters are computed from resource metrics.

Measurement directives - These specify how to measure a metric

Service Level Agreement Cont'd

There are two types of SLAs from the perspective of application hosting as described in detail here.

Infrastructure SLA.

- The infrastructure provider manages and offers guarantees on availability of the infrastructure, namely, server machine, power, network connectivity, and so on.
- Enterprises manage themselves, their applications that are deployed on these server machines.
- The machines are leased to the customers and are isolated from machines of other customers.
- In such dedicated hosting environments, a practical example of service-level guarantees offered by infrastructure providers.

Service Level Agreement Cont'd

Application SLA.

- In the application co-location hosting model, the server capacity is available to the applications based solely on their resource demands.
- Hence, the service providers are flexible in allocating and de-allocating computing resources among the co-located applications.

LIFE CYCLE OF SLA

Each SLA goes through a sequence of steps starting from identification of terms and conditions, activation and monitoring of the stated terms and conditions, and eventual termination of contract once the hosting relationship ceases to exist.

Service Level Agreement Cont'd

Such a sequence of steps is called SLA life cycle and consists of the following five phases:

1. Contract definition
2. Publishing and discovery
3. Negotiation
4. Operationalization
5. De-commissioning

Here, we explain in detail each of these phases of SLA life cycle.

Contract Definition.

Generally, service providers define a set of service offerings and corresponding SLAs using standard templates. These service offerings form a catalog. Individual SLAs for enterprises can be derived by customizing these base SLA templates.

Service Level Agreement Cont'd

Publication and Discovery.

Service provider advertises these base service offerings through standard publication media, and the customers should be able to locate the service provider by searching the catalog. The customers can search different competitive offerings and shortlist a few that fulfill their requirements for further negotiation.

Negotiation.

Once the customer has discovered a service provider who can meet their application hosting need, the SLA terms and conditions needs to be mutually agreed upon before signing the agreement for hosting the application. For a standard packaged application which is offered as service, this phase could be automated.

Service Level Agreement Cont'd

For customized applications that are hosted on cloud platforms, this phase is manual. The service provider needs to analyze the application's behavior with respect to scalability and performance before agreeing on the specification of SLA. At the end of this phase, the SLA is mutually agreed by both customer and provider and is eventually signed off. SLA negotiation can utilize the WS-negotiation specification.

Operationalization.

SLA operation consists of SLA monitoring, SLA ac- counting, and SLA enforcement. SLA monitoring involves measuring parameter values and calculating the metrics defined as a part of SLA and determining the deviations.

Service Level Agreement Cont'd

On identifying the deviations, the concerned parties are notified.

SLA accounting involves capturing and archiving the SLA adherence for compliance. As part of accounting, the application's actual performance and the performance guaranteed as a part of SLA is reported.

De-commissioning.

SLA decommissioning involves termination of all activities performed under a particular SLA when the hosting relationship between the service provider and the service consumer has ended. SLA specifies the terms and conditions of contract termination and specifies situations under which the relationship between a service provider and a service consumer can be considered to be legally ended.

Service Level Agreement Cont'd

SLA MANAGEMENT IN CLOUD

SLA management of applications hosted on cloud platforms involves five phases.

1. Feasibility 2. On-boarding 3. Pre-production
4. Production 5. Termination

1. Feasibility Analysis

MSP conducts the feasibility study of hosting an application on their cloud platforms. This study involves three kinds of feasibility:

(1) technical feasibility,
(2) Infra structure feasibility, and
(3) financial feasibility.

Service Level Agreement Cont'd

Technical feasibility of an application implies determining the following:

1. Ability of an application to scale out.
2. Compatibility of the application with the cloud platform being used within the MSP's data center.
3. The need and availability of a specific hardware and software required for hosting and running of the application.
4. Preliminary information about the application performance and whether be met by the MSP.

Infrastructure feasibility involves determining the availability of infrastructural resources in sufficient quantity so that the projected demands of the application can be met

Service Level Agreement Cont'd

Financial feasibility study involves determining the approximate cost to be incurred by the MSP and the price the MSP charges the customer so that the hosting activity is profitable to both of them A feasibility report consist s of the results of the above three feasibility studies. The report forms the basis for further communication with the customer . Once the provider and customer agree upon the findings of the report , the outsourcing of the application hosting activity proceeds to the next phase, called "onboarding" of application.

Only the basic feasibility of hosting an application has been carried in this phase. However , the detailed runtime characteristics of the application are studied as par t of the on- boarding activity.

Service Level Agreement Cont'd

2. On-Boarding of Application

Once the customer and the MSP agree in principle to host the application based on the findings of the feasibility study, the application is moved from the customer servers to the hosting platform. The application is accessible to its end users only after the on- boarding activity is completed.

On-boarding activity consists of the following steps:

a) Packing of the application for deploying on physical or virtual environments. Application packaging is the process of creating deployable components on the hosting platform (could be physical or virtual). Open Virtualization Format (OVF) standard is used for packaging the application for cloud platform .

Service Level Agreement Cont'd

b) The packaged application is executed directly on the physical servers to capture and analyze the application performance characteristics.

c) The application is executed on a virtualized platform and the application performance characteristics are noted again.

d) Based on the measured performance characteristics, different possible SLAs are identified. The resources required and the costs involved for each SLA are also computed.

e) Once the customer agrees to the set of SLOs and the cost, the MSP starts creating different policies required by the data center for automated management of the application. These policies are of three types: (1) business, (2) operational, and (3) provisioning.

Service Level Agreement Cont'd

3. Preproduction

Once the determination of policies is completed as discussed in previous phase, the application is hosted in a simulated production environment. Once both parties agree on the cost and the terms and conditions of the SLA, the customer sign-off is obtained. On successful completion of this phase the MSP allows the application to go on-live.

4. Production

In this phase, the application is made accessible to its end users under the agreed SLA. In the case of the former, on-boarding activity is repeated to analyze the application and its policies with respect to SLA fulfillment.

Service Level Agreement Cont'd

In case of the latter, a new set of policies are formulated to meet the fresh terms and conditions of the SLA.

5. Termination

When the customer wishes to withdraw the hosted application and does not wish to continue to avail the services of the MSP for managing the hosting of its application, the termination activity is initiated.

Business Benefits of Cloud Computing

There are some clear business benefits to building applications in the cloud. A few of these are listed here:

Almost Zero Upfront Infrastructure Investment.

If you have to build a large-scale system, it may cost a fortune to invest in real estate, physical security, hardware (racks, servers, routers, backup power supplies), hardware management (power management, cooling), and operations personnel.

Because of the high upfront costs, the project would typically require several rounds of management approvals before the project could even get started.

Now, with utility-style cloud computing, there is no fixed cost or startup cost.

Business Benefits of Cloud Computing Cont'd

Just-in-Time Infrastructure.

In the past, if your application became popular and your systems or your infrastructure did not scale, you became a victim of your own success. Conversely, if you invested heavily and did not get popular, you became a victim of your failure.

By deploying applications in-the-cloud with just-in-time self-provisioning, you do not have to worry about pre-procuring capacity for large-scale systems.

This increases agility, lowers risk, and lowers operational cost because you scale only as you grow and only pay for what you use.

Business Benefits of Cloud Computing Cont'd

More Efficient Resource Utilization.

System administrators usually worry about procuring hardware (when they run out of capacity) and higher infrastructure utilization (when they have excess and idle capacity).

With the cloud, they can manage resources more effectively and efficiently by having the applications request and relinquish resources on-demand.

Usage-Based Costing.

With utility-style pricing, you are billed only for the infrastructure that has been used. You are not paying for allocated infrastructure but instead for unused infrastructure. This adds a new dimension to cost savings.

Business Benefits of Cloud Computing Cont'd

You can see immediate cost savings (some- times as early as your next month's bill) when you deploy an optimization patch to update your cloud application.

For example, if a caching layer can reduce your data requests by 70%, the savings begin to accrue immediately and you see the reward right in the next bill. Moreover, if you are building platforms on the top of the cloud, you can pass on the same flexible, variable usage-based cost structure to your own customers

Reduced Time to Market.

Parallelization is one of the great ways to speed up processing. If one compute-intensive or data-intensive job that can be run in parallel takes 500 hours to process on one machine, with cloud architectures ,

Business Benefits of Cloud Computing Cont'd

it would be possible to spawn and launch 500 instances and process the same job in 1 hour.

Having available an elastic infrastructure provides the application with the ability to exploit parallelization in a cost-effective manner reducing time to market.

.

Technical Benefits of Cloud Computing

Some of the technical benefits of cloud computing includes:

Automation "Scriptable Infrastructure":

You can create repeatable build and deployment systems by leveraging programmable (API-driven) infrastructure

Auto-scaling:

You can scale your applications up and down to match your unexpected demand without any human intervention. Auto-scaling encourages automation and drives more efficiency.

Proactive Scaling:

Scale your application up and down to meet your anticipated demand with proper planning understanding of your traffic patterns so that you keep your costs low while scaling.

Technical Benefits of Cloud Computing Cont'd

More Efficient Development Life Cycle:

Production systems may be easily cloned for use as development and test environments. Staging environments may be easily promoted to production.

Improved Testability:

Never run out of hardware for testing. Inject and automate testing at every stage during the development process. You can spawn up an instant test lab with preconfigured environments only for the duration of testing phase.

Disaster Recovery and Business Continuity:

The cloud provides a lower cost option for maintaining a fleet of DR servers and data storage.

Technical Benefits of Cloud Computing Cont'd

With the cloud, you can take advantage of geo-distribution and replicate the environment in other location within minutes. "Overflow" the Traffic to the Cloud:

With a few clicks and effective load balancing tactics, you can create a complete overflow-proof application by routing excess traffic to the cloud.

Cloud Standardization

Cloud Standardization

At its most basic, cloud computing is simply the delivery of applications; security and other services; storage and other infrastructures; and platforms such as those for software development to users over the Internet or a private cloud.

Cloud computing appeals to many organizations because it minimizes the amount of hardware and software that users must own, maintain, and upgrade. In essence, users pay only for the computing capability they need.

Standardization issues

Many of today's in-progress standards, are being based in part on the US National Institute of Standards and Technology's Special Publication 800-145, a document called "The NIST Definition of Cloud Computing (Draft).

Cloud Standardization Cont'd

Table 1. Comparison of cloud-computing standards.

Organization	Working Group	Standard	Purpose
Distributed Management Task Force		Open Virtualization Format (OVF)	Establishes a transport mechanism for moving virtual machines from one hosted platform to another
IEEE	P2301 P2302	P2301: Guide for Cloud Portability and Interoperability Profiles (CPIP) P2302: Standard for Intercloud Interoperability and Federation (SIIF)	CPIP: Metastandard with profiles for existing and in-progress cloud computing standards in areas such as applications, portability, and management. SIIF: Establishes the characteristics necessary to create cloud interoperability and federation.
Open Grid Forum	Open Cloud Computing Interface	Open Cloud Computing Interface (OCCI)	Develop APIs for cloud management tasks. APIs enable interfacing between IaaS cloud implementations.
Organization for the Advancement of Structured Information Standards (OASIS)	IDCCloud Technical Committee Symptoms Automation Framework Technical Committee		ID Cloud focuses on security issues such as identity management and vulnerability mitigation. Symptoms Automation Framework establishes communications so that cloud providers understand consumer requirements.
Storage Networking Industry Association	Cloud Storage Initiative	Cloud Data Management Interface (CDMI)	Provides standardization for client interactions with cloud-based storage, cloud data management, and cloud-to-cloud storage interactions.

Cloud Computing Management and Application Trends Cont'd

These 5 innovative trends in Cloud Computing can benefit small to large businesses altogether:

1. The Emergence of Digital Natives in the Workforce
2. Artificial Intelligence (AI)
3. Hybrid Cloud Computing
4. Quantum Computing
5. Serverless Computing

The End.

Thank You.

Q & A