ISM V Notes - dereer

Information Security (SRM Institute of Science and Technology)

Chapter 13
Cloud Computing

KEY CONCEPTS

- Essential Characteristics of Cloud Computing
- Cloud Services and Deployment Models
- Cloud Computing Infrastructure
- Cloud Adoption Considerations

In today's competitive environment, organizations are under increasing pressure to improve efficiency and transform their IT processes to achieve more with less. Businesses need reduced time-to-market, better agility, higher availability, and reduced expenditures to meet the changing business requirements and accelerated pace of innovation. These business requirements are posing several challenges to IT teams. Some of the key challenges are serving customers worldwide around the clock, refreshing technology quickly and faster provisioning of IT resources — all at reduced costs.

These long-standing challenges are addressed with the emergence of a new computing style, called *cloud computing*, which enables organizations and individuals to obtain and provision IT resources as a service. With cloud computing, users can browse and select relevant cloud services, such as compute, software, storage, or a combination of these resources, via a portal. Cloud computing automates delivery of selected cloud services to the users. It helps organizations and individuals deploy IT resources at reduced total cost of ownership with faster provisioning and compliance adherence.

A widely adopted definition of cloud computing comes from the U.S. National Institute of Standards and Technology (NIST Special Publication 800-145):

> *Cloud computing is a model for enabling ubiquitous, 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 chapter covers the enabling technologies, essential characteristics, benefits, services, deployment models, and infrastructure of cloud computing. The chapter also includes the challenges and considerations in adopting cloud computing.

13.1 Cloud Enabling Technologies

Grid computing, utility computing, virtualization, and service-oriented architecture are enabling technologies of cloud computing.

- *Grid computing* is a form of distributed computing that enables the resources of numerous heterogeneous computers in a network to work together on a single task at the same time. Grid computing enables parallel computing and is best for large workloads.
- *Utility computing* is a service-provisioning model in which a service provider makes computing resources available to customers, as required, and charges them based on usage. This is analogous to other utility services, such as electricity, where charges are based on the consumption.
- *Virtualization* is a technique that abstracts the physical characteristics of IT resources from resource users. It enables the resources to be viewed and managed as a pool and lets users create virtual resources from the pool. Virtualization provides better flexibility for provisioning of IT resources compared to provisioning in a non-virtualized environment. It helps optimize resource utilization and delivering resources more efficiently.
- *Service Oriented Architecture* (SOA) provides a set of services that can communicate with each other. These services work together to perform some activity or simply pass data among services.

13.2 Characteristics of Cloud Computing

A computing infrastructure used for cloud services must have certain capabilities or characteristics. According to NIST, the cloud infrastructure should have five essential characteristics:

- **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 provider.

  A cloud service provider publishes a service catalogue, which contains information about all cloud services available to consumers. The service catalogue includes information about service attributes, prices, and request processes. Consumers view the service catalogue via a web-based user interface and use it to request for a service. Consumers can either leverage the "ready-to-use" services or change a few service parameters to customize the services.

- **Broad network access:** Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (for example, mobile phones, tablets, laptops, and workstations).
- **Resource pooling:** The provider's computing resources are pooled to serve multiple consumers using a multitenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. 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 (for example, country, state, or data center). Examples of resources include storage, processing, memory, and network bandwidth.
- **Rapid elasticity:** Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

  Consumers can leverage rapid elasticity of the cloud when they have a fluctuation in their IT resource requirements. For example, an organization might require double the number of web and application servers for a specific duration to accomplish a specific task. For the remaining period, they might want to release idle server resources to cut down the expenses. The cloud enables consumers to grow and shrink the demand for resources dynamically.
- **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 (for example, storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

MULTITENANCY

Multitenancy **refers to an architecture in which multiple independent consumers (tenants) are serviced using a single set of resources. This lowers the cost of services for consumers. Virtualization enables resource pooling and multitenancy in the cloud. For example, multiple virtual machines from different consumers can run simultaneously on the same physical server that runs the hypervisor.**

13.3 Benefits of Cloud Computing

Cloud computing offers the following key benefits:

- **Reduced IT cost:** Cloud services can be purchased based on pay-per-use or subscription pricing. This reduces or eliminates the consumer's IT capital expenditure (CAPEX).
- **Business agility:** Cloud computing provides the capability to allocate and scale computing capacity quickly. Cloud computing can reduce the time required to provision and deploy new applications and services from months to minutes. This enables businesses to respond more quickly to market changes and reduce time-to-market.
- **Flexible scaling:** Cloud computing enables consumers to scale up, scale down, scale out, or scale in the demand for computing resources easily. Consumers can unilaterally and automatically scale computing resources without any interaction with cloud service providers. The flexible service provisioning capability of cloud computing often provides a sense of unlimited scalability to the cloud service consumers.
- **High availability:** Cloud computing has the capability to ensure resource availability at varying levels depending on the consumer's policy and priority. Redundant infrastructure components (servers, network paths, and storage equipment, along with clustered software) enable fault tolerance for cloud deployments. These techniques can encompass multiple data centers located in different geographic regions, which prevents data unavailability due to regional failures.

13.4 Cloud Service Models

According to NIST, cloud service offerings are classified primarily into three models: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS).

13.4.1 Infrastructure-as-a-Service

The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems and deployed applications; and possibly limited control of select networking components (for example, host firewalls).

IaaS is the base layer of the cloud services stack (see Figure 13-1 [a]). It serves as the foundation for both the SaaS and PaaS layers.

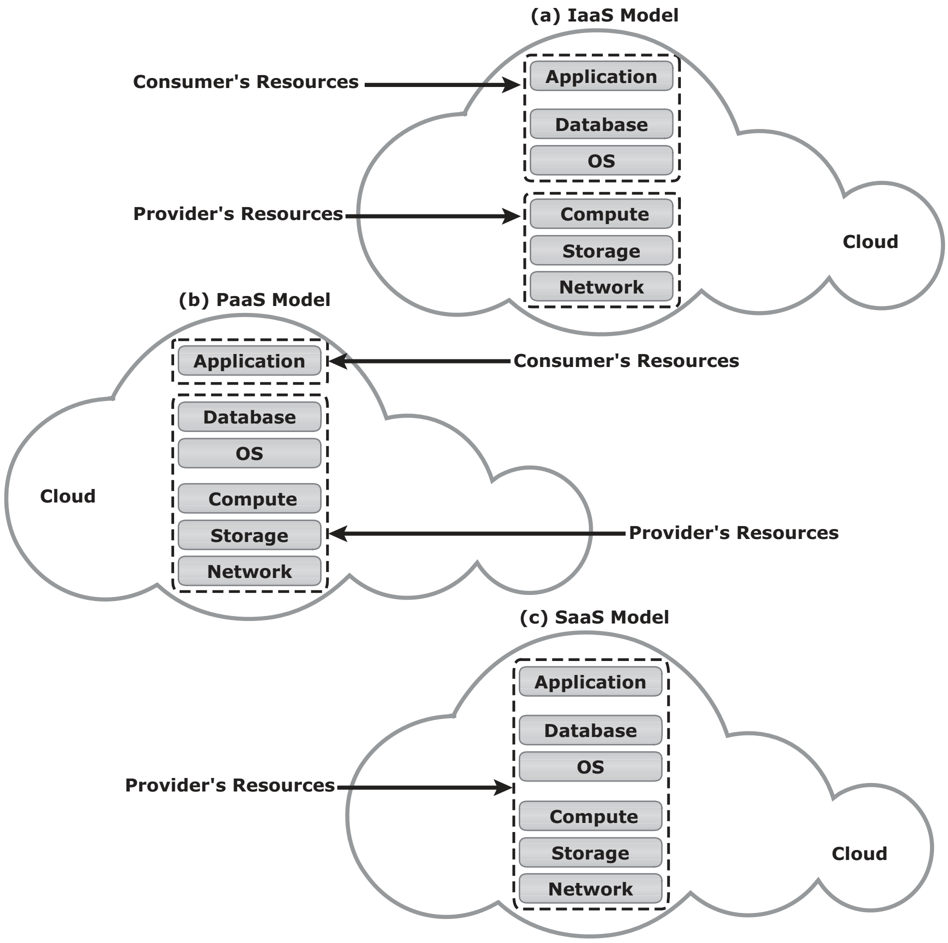


Figure 13-1: IaaS, PaaS, and SaaS models

Amazon Elastic Compute Cloud (Amazon EC2) is an example of IaaS that provides scalable compute capacity, on-demand, in the cloud. It enables consumers to leverage Amazon's massive computing infrastructure with no up-front capital investment.

13.4.2 Platform-as-a-Service

The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment. (See Figure 13-1 [b]).

PaaS is also used as an application development environment, offered as a service by the cloud service provider. The consumer may use these platforms to code their applications and then deploy the applications on the cloud. Because the workload to the deployed applications varies, the scalability of computing resources is usually guaranteed by the computing platform, transparently. Google App Engine and Microsoft Windows Azure Platform are examples of PaaS.

13.4.3 Software-as-a-Service

The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (for example, web-based e-mail), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings. (See Figure 13-1[c]).

In a SaaS model, applications, such as customer relationship management (CRM), e-mail, and instant messaging (IM), are offered as a service by the cloud service providers. The cloud service providers exclusively manage the required computing infrastructure and software to support these services. The consumers may be allowed to change a few application configuration settings to customize the applications.

EMC Mozy is an example of SaaS. Consumers can leverage the Mozy console to perform automatic, secured, online backup and recovery of their data with ease. Salesforce.com is a provider of SaaS-based CRM applications, such as Sales Cloud and Service Cloud.

13.5 Cloud Deployment Models

According to NIST, cloud computing is classified into four deployment models — public, private, community, and hybrid — which provide the basis for how cloud infrastructures are constructed and consumed.

13.5.1 Public Cloud

In a *public cloud* model, the cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

Consumers use the cloud services offered by the providers via the Internet and pay metered usage charges or subscription fees. An advantage of the public cloud is its low capital cost with enormous scalability. However, for consumers, these benefits come with certain risks: no control over the resources in the cloud, the security of confidential data, network performance, and interoperability issues. Popular public cloud service providers are Amazon, Google, and Salesforce.com. Figure 13-2 shows a public cloud that provides cloud services to organizations and individuals.

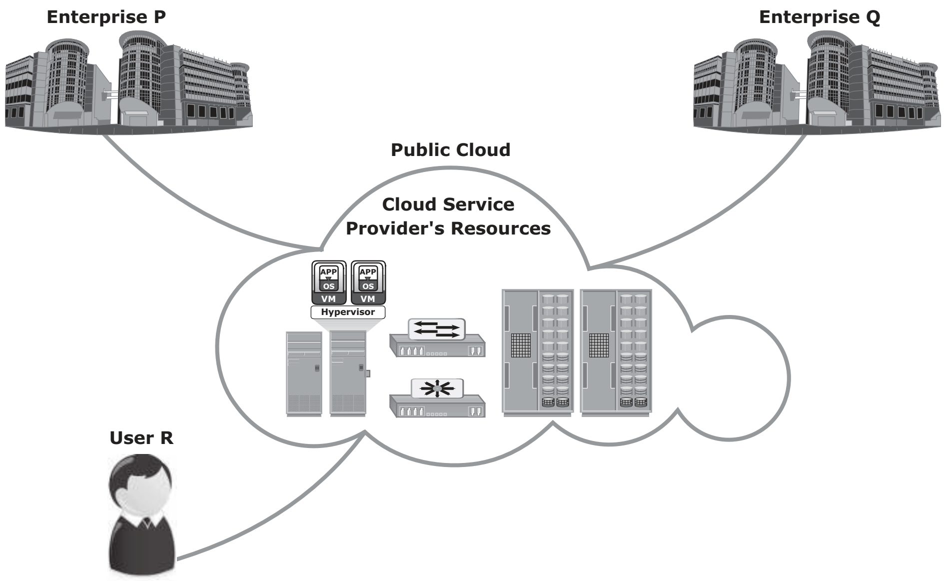
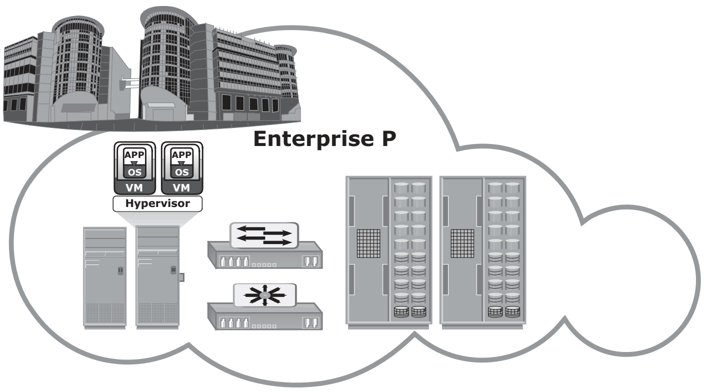


Figure 13-2: Public cloud

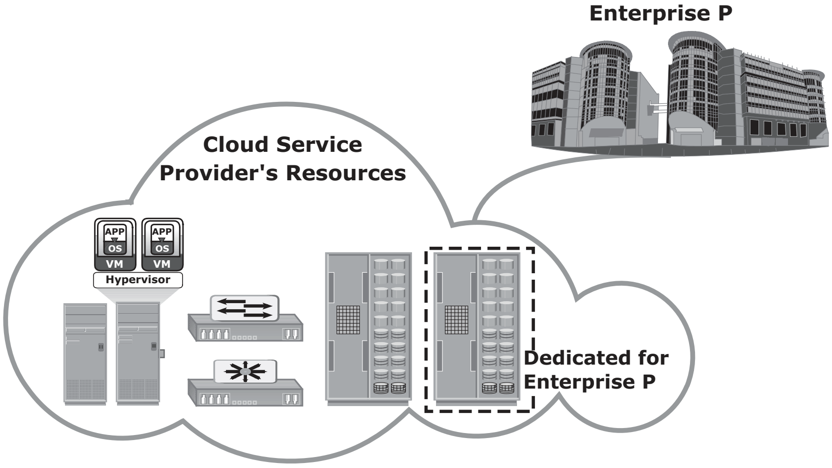
13.5.2 Private Cloud

In a *private cloud* model, the cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (for example, business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises. Following are two variations to the private cloud model:

- **On-premise private cloud:** The on-premise private cloud, also known as internal cloud, is hosted by an organization within its own data centers (see Figure 13-3 [a]). This model enables organizations to standardize their cloud service management processes and security, although this model has limitations in terms of size and resource scalability. Organizations would also need to incur the capital and operational costs for the physical resources. This is best suited for organizations that require complete control over their applications, infrastructure configurations, and security mechanisms.



(a) On-Premise Private Cloud



(b) Externally Hosted Private Cloud

Figure 13-3: On-premise and externally hosted private clouds

- **Externally hosted private cloud:** This type of private cloud is hosted external to an organization (see Figure 13-3 [b]) and is managed by a third-party organization. The third-party organization facilitates an exclusive cloud environment for a specific organization with full guarantee of privacy and confidentiality.

13.5.3 Community Cloud

In a *community cloud* model, the cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (for example, mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises. (See Figure 13-4).

In a community cloud, the costs spread over to fewer consumers than a public cloud. Hence, this option is more expensive but might offer a higher level of privacy, security, and compliance. The community cloud also offers organizations access to a vast pool of resources compared to the private cloud. An example in which a community cloud could be useful is government agencies. If various agencies within the government operate under similar guidelines, they could all share the same infrastructure and lower their individual agency's investment.

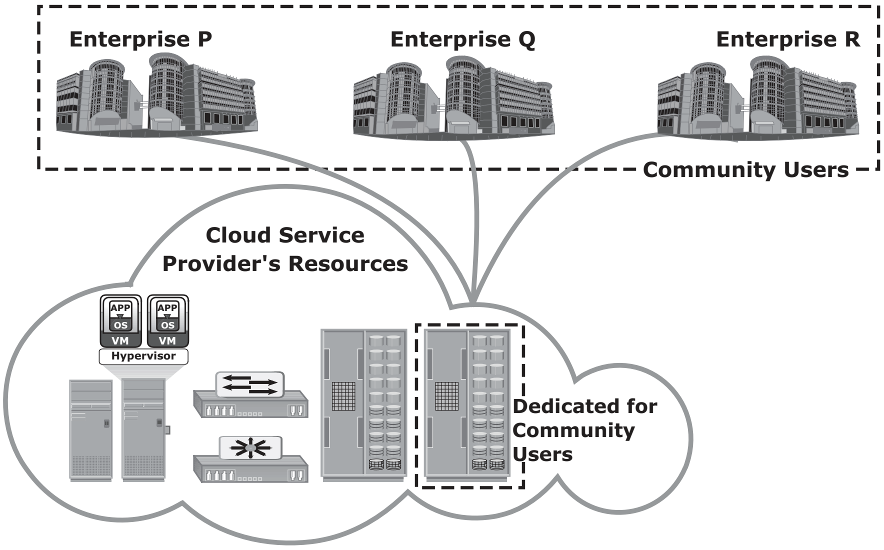


Figure 13-4: Community cloud

13.5.4 Hybrid Cloud

In a *hybrid cloud* model, the cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (for example, cloud bursting for load balancing between clouds).

The hybrid model allows an organization to deploy less critical applications and data to the public cloud, leveraging the scalability and cost-effectiveness of the public cloud. The organization's mission-critical applications and data remain on the private cloud that provides greater security. Figure 13-5 shows an example of a hybrid cloud.

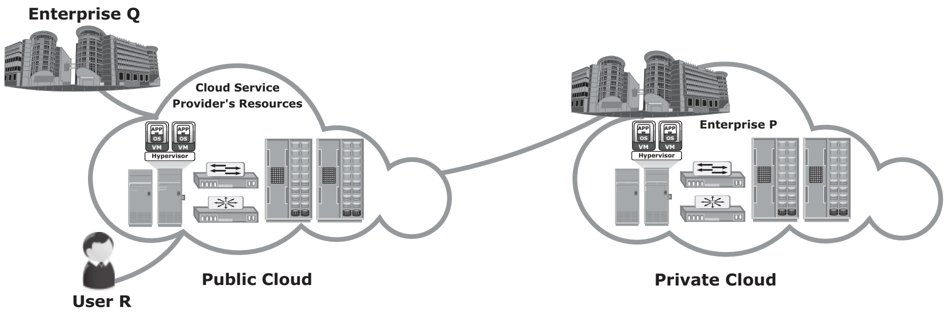


Figure 13-5: Hybrid cloud

13.6 Cloud Computing Infrastructure

A cloud computing infrastructure is the collection of hardware and software that enables the five essential characteristics of cloud computing. Cloud computing infrastructure usually consists of the following layers:

- Physical infrastructure
- Virtual infrastructure
- Applications and platform software
- Cloud management and service creation tools

The resources of these layers are aggregated and coordinated to provide cloud services to the consumers (see Figure 13-6).

13.6.1 Physical Infrastructure

The physical infrastructure consists of physical computing resources, which include physical servers, storage systems, and networks. Physical servers are connected to each other, to the storage systems, and to the clients via networks, such as IP, FC SAN, IP SAN, or FCoE networks.

Cloud service providers may use physical computing resources from one or more data centers to provide services. If the computing resources are distributed across multiple data centers, connectivity must be established among them. The connectivity enables the data centers in different locations to work as a single large data center. This enables migration of business applications and data across data centers and provisioning cloud services using the resources from multiple data centers.

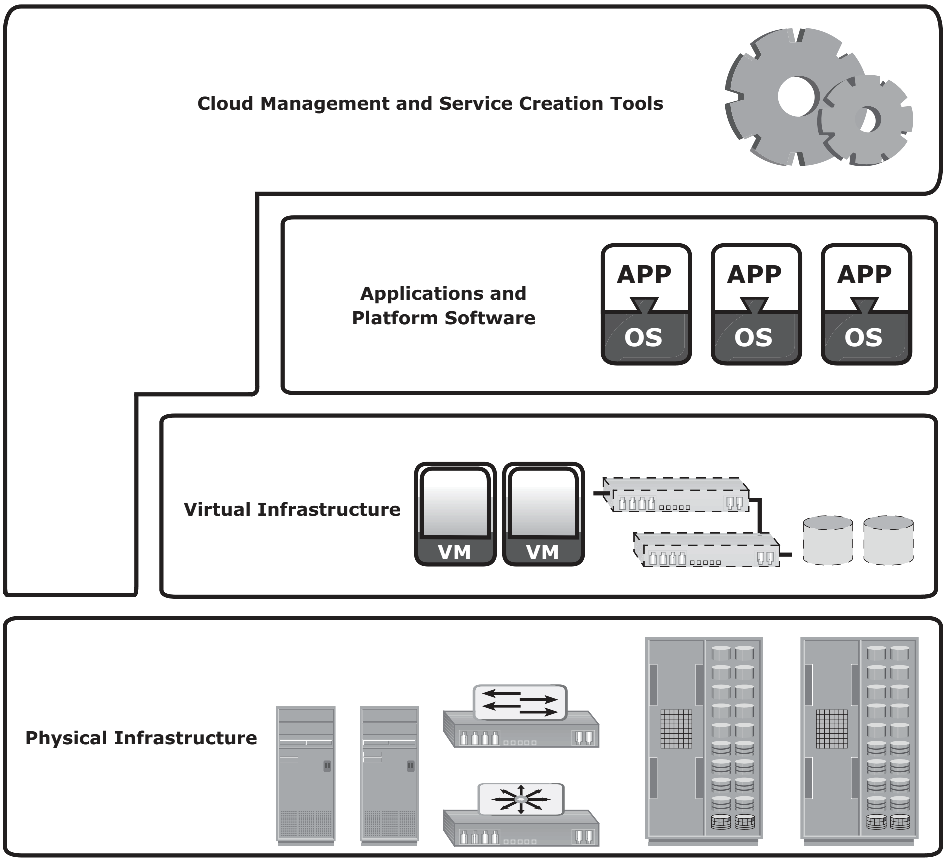


Figure 13-6: Cloud infrastructure layers

13.6.2 Virtual Infrastructure

Cloud service providers employ virtualization technologies to build a virtual infrastructure layer on the top of the physical infrastructure. Virtualization enables fulfilling some of the cloud characteristics, such as resource pooling and rapid elasticity. It also helps reduce the cost of providing the cloud services. Some cloud service providers may not have completely virtualized their physical infrastructure yet, but they are adopting virtualization for better efficiency and optimization.

Virtualization abstracts physical computing resources and provides a consolidated view of the resource capacity. The consolidated resources are managed as a single entity called a *resource pool*. For example, a resource pool might group CPUs of physical servers within a cluster. The capacity of the resource pool is the sum of the power of all CPUs (for example, 10,000 megahertz) available in the cluster. In addition to the CPU pool, the virtual infrastructure includes other types of resource pools, such as memory pool, network pool, and storage pool. Apart from resource pools, the virtual infrastructure also includes *identity pools*, such as VLAN ID pools and VSAN ID pools. The number of each type of pool and the pool capacity depend on the cloud service provider's requirement to create different cloud services.

Virtual infrastructure also includes virtual computing resources, such as virtual machines, virtual storage volumes, and virtual networks. These resources obtain capacities, such as CPU power, memory, network bandwidth, and storage space from the resource pools. The capacity is allocated to the virtual computing resources easily and flexibly based on the service requirement. Virtual networks are created using network identifiers, such as VLAN IDs and VSAN IDs from the respective identity pools. Virtual computing resources are used for creating cloud infrastructure services.

13.6.3 Applications and Platform Software

This layer includes a suite of business applications and platform software, such as the OS and database. Platform software provides the environment on which business applications run. Applications and platform software are hosted on virtual machines to create SaaS and PaaS. For SaaS, both the application and platform software are provided by cloud service providers. In the case of PaaS, only the platform software is provided by cloud service providers; consumers export their applications to the cloud.

13.6.4 Cloud Management and Service Creation Tools

The cloud management and service creation tools layer includes three types of software:

- Physical and virtual infrastructure management software
- Unified management software
- User-access management software

This classification is based on the different functions performed by the software. This software interacts with each other to automate provisioning of cloud services.

The physical and virtual infrastructure management software is offered by the vendors of various infrastructure resources and third-party organizations. For example, a storage array has its own management software. Similarly, network and physical servers are managed independently using network and compute management software respectively. This software provides interfaces to construct a virtual infrastructure from the underlying physical infrastructure.

Unified management software interacts with all standalone physical and virtual infrastructure management software. It collects information on the existing physical and virtual infrastructure configurations, connectivity, and utilization. Unified management software compiles this information and provides a consolidated view of infrastructure resources scattered across one or more data centers. It allows an administrator to monitor performance, capacity, and availability of physical and virtual resources centrally. Unified management software also provides a single management interface to configure physical and virtual infrastructure and integrate the compute (both CPU and memory), network, and storage pools. The integration allows a group of compute pools to use the storage and network pools for storing and transferring data respectively. The unified management software passes configuration commands to respective physical and virtual infrastructure management software, which executes the instructions. This eliminates the administration of compute, storage, and network resources separately using native management software.

The key function of the unified management software is to automate the creation of cloud services. It enables administrators to define service attributes such as CPU power, memory, network bandwidth, storage capacity, name and description of applications and platform software, resource location, and backup policy. When the unified management software receives consumer requests for cloud services, it creates the service based on predefined service attributes.

The user-access management software provides a web-based user interface to consumers. Consumers can use the interface to browse the service catalogue and request cloud services. The user-access management software authenticates users before forwarding their request to the unified management software. It also monitors allocation or usage of resources associated to the cloud service instances. Based on the allocation or usage of resources, it generates a chargeback report. The chargeback report is visible to consumers and provides transparency between consumers and providers.

CLOUD-OPTIMIZED STORAGE

Content-rich applications combined with the growth of user-generated unstructured data is challenging to manage with the traditional approach of storing data at scale. This combination of massive growth, new information types, and the need to serve multiple locations and users around the world, has led to requirements for information storage and management at a global scale. Cloud-optimized storage is a solution to meet these requirements. It delivers scalable and flexible architecture that provides rapid elasticity, global access, and storage capacity on-demand. It also addresses the constraints of rigid, mount-point based interaction between storage and consumer by presenting a singular access point to the entire storage infrastructure.

(Continued)

CLOUD-OPTIMIZED STORAGE *(continued)*

It leverages a built-in multitenancy model and enables self-service; fully metered access to storage resources thereby delivers storage-as-a-service on a shared infrastructure. Cloud-optimized storage typically leverages object-based storage technology that uses customizable, value-driven metadata to drive storage placement, protection, and life cycle policies. Following are key characteristics of cloud-optimized storage solution:

- **Massively scalable infrastructure that supports a large number of objects across a globally distributed infrastructure**
- **Unified namespace that eliminates capacity, location, and other file system limitations**
- **Metadata and policy-based information management capabilities that optimize data protection, availability, and cost, based on service levels**
- **Secure multitenancy that enables multiple applications to be securely served from the same infrastructure. Each application is securely partitioned and data is neither co-mingled nor accessible by other tenants.**
- **Provides access through REST and SOAP web service APIs and file-based access using a variety of client devices**

13.7 Cloud Challenges

Although there is growing acceptance of cloud computing, both the cloud service consumers and providers have been facing some challenges.

13.7.1 Challenges for Consumers

Business-critical data requires protection and continuous monitoring of its access. If the data moves to a cloud model other than an on-premise private cloud, consumers could lose absolute control of their sensitive data. Although most of the cloud service providers offer enhanced data security, consumers might not be willing to transfer control of their business-critical data to the cloud.

Cloud service providers might use multiple data centers located in different countries to provide cloud services. They might replicate or move data across these data centers to ensure high availability and load distribution. Consumers may or may not know in which country their data is stored. Some cloud service providers allow consumers to select the location for storing their data. Data privacy concerns and regulatory compliance requirements, such as the EU Data Protection Directive and the U.S. Safe Harbor program, create challenges for the consumers in adopting cloud computing.

Cloud services can be accessed from anywhere via a network. However, network latency increases when the cloud infrastructure is not close to the access point. A high network latency can either increase the application response time or cause the application to timeout. This can be addressed by implementing stringent Service Level Agreements (SLAs) with the cloud service providers.

Another challenge is that cloud platform services may not support consumers' desired applications. For example, a service provider might not be able to support highly specialized or proprietary environments, such as compatible OSs and preferred programming languages, required to develop and run the consumer's application. Also, a mismatch between hypervisors could impact migration of virtual machines into or between clouds.

Another challenge is vendor lock-in: the difficulty for consumers to change their cloud service provider. A lack of interoperability between the APIs of different cloud service providers could also create complexity and high migration costs when moving from one service provider to another.

13.7.2 Challenges for Providers

Cloud service providers usually publish a service-level agreement (SLA) so that their consumers know about the availability of service, quality of service, downtime compensation, and legal and regulatory clauses. Alternatively, customer-specific SLAs may be signed between a cloud service provider and a consumer. SLAs typically mention a penalty amount if cloud service providers fail to provide the service levels. Therefore, cloud service providers must ensure that they have adequate resources to provide the required levels of services. Because the cloud resources are distributed and service demands fluctuate, it is a challenge for cloud service providers to provision physical resources for peak demand of all consumers and estimate the actual cost of providing the services.

Many software vendors do not have a cloud-ready software licensing model. Some of the software vendors offer standardized cloud licenses at a higher price compared to traditional licensing models. The cloud software licensing complexity has been causing challenges in deploying vendor software in the cloud. This is also a challenge to the consumer.

Cloud service providers usually offer proprietary APIs to access their cloud. However, consumers might want open APIs or standard APIs to become the tenant of multiple clouds. This is a challenge for cloud service providers because this requires agreement among cloud service providers.

13.8 Cloud Adoption Considerations

Organizations that decide to adopt cloud computing always face this question: "How does the cloud fit the organization's environment?" Most organizations are not ready to abandon their existing IT investments to move all their business processes to the cloud at once. Instead, they need to consider various factors before moving their business processes to the cloud. Even individuals seeking to use cloud services need to understand some cloud adoption considerations. Following are some key considerations for cloud adoption:

- **Selection of a deployment model:** Risk versus convenience is a key consideration for deciding on a cloud adoption strategy. This consideration also forms the basis for choosing the right cloud deployment model. A public cloud is usually preferred by individuals and start-up businesses. For them, the cost reduction offered by the public cloud outweighs the security or availability risks in the cloud. Small- and medium-sized businesses (SMBs) have a moderate customer base, and any anomaly in customer data and service levels might impact their business. Therefore, they may not be willing to deploy their tier 1 applications, such as Online Transaction Processing (OLTP), in the public cloud. A hybrid cloud model fits in this case. The tier 1applications should run on the private cloud, whereas less critical applications such as backup, archive, and testing can be deployed in the public cloud. Enterprises typically have a strong customer base worldwide. They usually enforce strict security policies to safeguard critical customer data. Because they are financially capable, they might prefer building their own private clouds.
- **Application suitability:** Not all applications are good candidates for a public cloud. This may be due to the incompatibility between the cloud platform software and the consumer applications, or maybe the organization plans to move a legacy application to the cloud. Proprietary and mission-critical applications are core and essential to the business. They are usually designed, developed, and maintained in-house. These applications often provide competitive advantages. Due to high security risk, organizations are unlikely to move these applications to the public cloud. These applications are good candidate for an on-premise private cloud. Nonproprietary and nonmission critical applications are suitable for deployment in the public cloud. If an application workload is network traffic-intensive, its performance might not be optimal if deployed in the public cloud. Also if the application communicates with other data center resources or applications, it might experience performance issues.
- **Financial advantage:** A careful analysis of financial benefits provides a clear picture about the cost-savings in adopting the cloud. The analysis should compare both the Total Cost of Ownership (TCO) and the Return on Investment (ROI) in the cloud and noncloud environment and identify the potential cost benefit. While calculating TCO and ROI, organizations and individuals should consider the expenditure to deploy and maintain their own infrastructure versus cloud-adoption costs. While calculating the expenditures for owning infrastructure resources, organizations should include both the capital expenditure (CAPEX) and operation expenditure

(OPEX). The CAPEX includes the cost of servers, storage, OS, application, network equipment, real estate, and so on. The OPEX includes the cost incurred for power and cooling, personnel, maintenance, backup, and so on. These expenditures should be compared with the operation cost incurred in adopting cloud computing. The cloud adoption cost includes the cost of migrating to the cloud, cost to ensure compliance and security, and usage or subscription fees. Moving applications to the cloud reduces CAPEX, except when the cloud is built on-premise.

- **Selection of a cloud service provider:** The selection of the provider is important for a public cloud. Consumers need to find out how long and how well the provider has been delivering the services. They also need to determine how easy it is to add or terminate cloud services with the service provider. The consumer should know how easy it is to move to another provider, when required. They must assess how the provider fulfills the security, legal, and privacy requirements. They should also check whether the provider offers good customer service support.
- **Service-level agreement (SLA):** Cloud service providers typically mention quality of service (QoS) attributes such as throughput and uptime, along with cloud services. The QoS attributes are generally part of an SLA, which is the service contract between the provider and the consumers. The SLA serves as the foundation for the expected level of service between the consumer and the provider. Before adopting the cloud services, consumers should check whether the QoS attributes meet their requirements.

13.9 Concepts in Practice: Vblock

Vblock is a completely integrated cloud infrastructure offering that includes compute, storage, network, and virtualization products. These products are provided by EMC, VMware, and Cisco, who have formed a coalition to deliver Vblocks.

Vblocks enable organizations to build virtualized data centers and cloud infrastructures. Vblocks are pre-architected, preconfigured, pretested and have defined performance and availability attributes. Rather than customers buying and assembling individual cloud infrastructure components, Vblock provides a validated cloud infrastructure solution and is factory-ready for deployment and production. This saves significant cost and deployment time.

EMC Unified Infrastructure Manager (UIM) is the unified management solution for Vblocks. UIM provides a single point of management for Vblocks and manages multiple Vblocks. With UIM, cloud infrastructure services can be provisioned automatically based on provisioning best practices.

For more information on Vblock, visit `www.emc.com`.

Summary

Cloud computing, although evolving, is gaining popularity because consumers see a potential cost reduction and service providers see an opportunity to provide new services. Cloud computing has enabled IT organizations and individuals to gain benefits, such as automated and rapid resource provisioning, flexibility, high availability, and faster time to market at a reduced total cost of ownership. Although there are concerns and challenges, the benefits of cloud computing are compelling enough to adopt it.

For organizations that own traditional data centers, cloud adoption is like a journey. The journey begins with the consolidation of computing resources including compute systems, storage, and networks using virtualization technologies. Followed by virtualization of resources, organizations need to take the next step of implementing unified cloud infrastructure management tools and come up with the services catalog. Implementing proper service-management processes is a key to align the delivery of cloud services to the expectations of businesses and consumers.

This chapter detailed cloud characteristics, benefits, services, deployment models, and infrastructure. It also covered cloud challenges and adoption considerations. The next chapter focuses on securing the storage infrastructure, which also includes storage security considerations in virtualized and cloud environments.

EXERCISES

1. **What are the essential characteristics of cloud computing?**
2. **How does cloud computing bring in business agility?**
3. **Research Service Oriented Architecture and its application in cloud computing.**
4. **Research cloud orchestration.**
5. **Research various considerations for selecting a public cloud service provider.**
6. **What are the costs that should be evaluated to determine the financial advantage of cloud?**