CHAPTER 2

CLOUD DEPLOYMENT MODELS

2.1 INTRODUCTION

Cloud computing is an emerging style of computing where applications, data, and resources are provided to users as services over the Web. The services provided may be available globally, always on, low in cost, 'on demand', massively scalable, 'pay-as-you-grow'. Consumers of a service need to care only about what the service does for them, and not on how it is implemented. Cloud computing is a technology that allows users to access software applications, store information, develop and test new software, create virtual servers, draw on disparate IT resources, and more – all over the Internet (or other broad network).

Cloud computing is a model-driven methodology that provides configurable computing resources such as servers, networks, storage, and applications as and when required with minimum efforts over the Internet services. Cloud also indicates essential characteristics, delivery models, and deployment models.

This chapter visualizes several models for cloud computing, including private clouds (where the deployment is within the organization's firewall) and public clouds (where the application services and data are hosted by a third party outside the firewall). Consistent data availability and security is a critical success factor for any cloud deployment. Businesses need to ensure that data is adequately protected and can be restored in a timely fashion following any disruption event.

Clouds need a datacenter, but the aim of cloud computing is to eliminate the need to think about datacenters. A datacenter is a facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (e.g., air conditioning, fire suppression), and security devices.

Datacenters are tied to locality, with specific components including redundant power supplies, redundant communications, environmental controls, security devices, etc.

Clouds are location-independent, providing abstracted versions of datacenter components that are not tied to a specific datacenter: virtual servers, virtual storage, virtual networking, etc. Reliability and redundancy comes from cloud providers using multiple datacenters, so clouds almost certainly span one or more datacenters, but themselves are not datacenters.

2.2 CLOUD CHARACTERISTICS

Cloud carries the basic infrastructure characteristics that are helpful to deploy cloud service in a fast and cost-effective way (Figure 2.1). The following characteristics set apart cloud from other computing techniques.

2.2.1 *On-Demand 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.

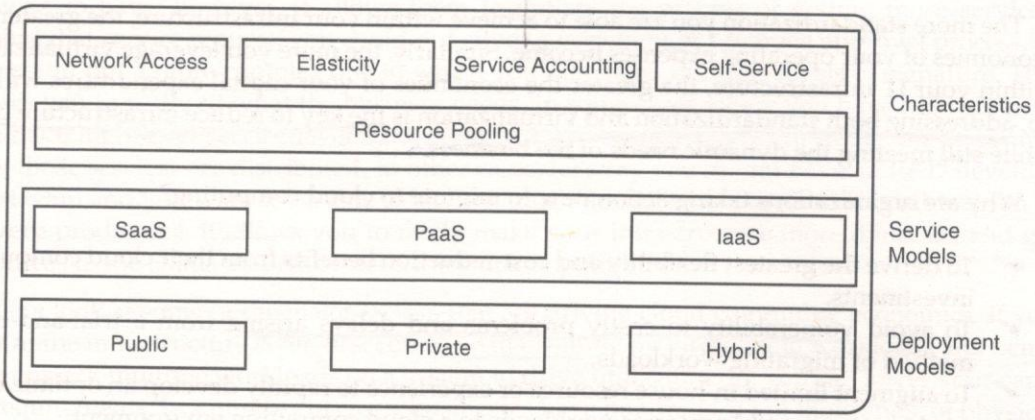

FIGURE 2.1 Cloud model.

2.2.2 *Ubiquitous Network Access*

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops and personal digital assistants [PDAs]).

2.2.3 *Location-Independent Resource Pooling (Multi-Tenant)*

The provider's computing resources are pooled to serve multiple customers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to the demand. There is a sense of location-independence in that the customer generally has no control or knowledge about the location where the services are located (for example, country, state, or datacenter). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

2.2.4 *Rapid Elasticity*

Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

2.3 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 (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the provider and the consumer of the utilized service.

The more standardization you are able to achieve within your infrastructure, the greater the economies of your operating expenses become. Similarly, the more you leverage virtualization within your IT infrastructure, the greater the economies of your capital expenditures will be. So, addressing both standardization and virtualization is the key to reduce infrastructure costs while still meeting the dynamic needs of the business.

Why are organizations taking action now to migrate to cloud computing?

- To derive the greatest flexibility and cost-reduction benefits from their cloud computing investments.
- To avoid vulnerability to costly problems and delays arising from a trial-and-error method of migrating workloads.
- To augment limited in-house resource or experience to rapidly develop an optimization roadmap to smoothly migrate workloads to a cloud computing environment.

Cloud vendors can address client's challenges by:

- Prioritizing workloads for cloud adoption based on business impact and risk.
- Maximizing business return by identifying applications that are well suited for cloud computing and have high business impact.
- Addressing problematic workloads to improve their propensity for cloud computing.
- Helping avoid costly implementation issues by identifying and addressing potential difficulties during migration.
- Mitigating the risk of costly implementation delays by identifying potential problems and addressing them before migration.
- Avoiding inadequate performance of highly complex and integrated workloads.
- Leveraging expertise to deliver an actionable roadmap to successfully migrate applications to a cloud computing environment.
- Accelerating your cloud initiatives.

2.3.1 *Cost Factor*

There are a number of reasons why cloud computing is popular with businesses. There is the cost aspect. By virtualizing your environment and standardizing it, you can deliver more services with fewer resources and drive up utilization. By adding automation, you can reduce labour cost giving you an additional cost benefit. This gives you a lot of flexibility because you can access cloud workloads services without thinking about the location and time of its execution. What this allows the organization to do then is to free up budget so that the money can be diverted to innovation and development of new capabilities rather than just keeping the lights on and running the IT enterprise.

The growing complexity of IT systems and soon a trillion connected things demand that sprawling processes become standardized services that are efficient, secure, and easy to access. A service management system will provide visibility, control, and automation across IT and business services to ensure consistent delivery. Self-service plus standardization will drive lower operational costs, unlock productivity, and ensure better security.

Cloud allows businesses to be smarter about how they deliver services. The first aspect of this is a self-service portal. This allows your end consumers to only see services they are

Self service portal

allowed to have; however, it allows them to initiate the process of getting those services. Behind that service request, you could put either a very light or no-touch approval process, or you could put a more complex one in which you may need multiple levels of signature. This allows you to really fit what the business needs. In some cases where you have high security, you have high-level service-level agreements (SLAs) – you really want to be able to control how those services are distributed. In other cases, let's say you do not have an R&D development team and you want to be able to have as much flexibility as possible. This allows you to be very productive. It allows you to really make your infrastructure more dynamic, and get resources to the teams that really need them at any point in time.

Let's look at some of the major factors that are driving cloud computing economics. If you look at the infrastructure layer, first comes virtualization. By virtualizing workloads and being able to stack multiple workloads on a system to drive utilization up, you can lower your capital requirements. In a number of cases, businesses have hundreds, if not thousands, of physical servers and unless they have used virtualization and unless they are really driving that utilization, the utilization could be as low as 10 percent. So, in a lot of cases, organizations that use cloud computing are able to drive utilization up and either lower future capital requirements or even retire antiquated equipment and drive their costs down.

From a labour perspective, using a self-service portal allows your clients to help themselves. So there is less support and it makes the offering more available from a service perspective. In terms of automation, it takes tasks that are very manual and repeatable, and by automation then, it reduces your IT operations cost. In a development or test environment, you need multiple skills to get that environment to the end-user. You need operating system skills, middleware skills, database skills, and application skills. This allows you to define that environment as a repeatable, deployable resource, and it drives down your labour cost there. Of course, you need to standardize those workloads. Standardization has labour cost and quality benefits so that you can ensure consistency from environment to environment.

In many cases, you may want to use multiple models for different types of services that you want to deliver. Starting with private cloud services, the first model (which is also the most popular currently) is the private on-premise cloud. If the cloud is within the organization's datacenter, it is operated and managed by the organization itself.

The need to achieving cost optimization has also provided fertile ground for cloud computing. The cloud paradigm is an attempt to improve service delivery by applying engineering discipline and economies of scale in an Internet-inspired architecture.

Cloud computing can be an important new option in helping businesses optimize the IT expense equation while maintaining fast, high-quality service delivery.

2.3.2 *Benefits*

We can enjoy many benefits by adopting cloud:

- **Self-service capability:** Once somebody deploys the cloud services, they are capable of self-service. Now testing teams do not have to buy computing services as they can enjoy the same services over the cloud and it reduces the procurement process. Hence, they can concentrate on the testing services and efforts.

- **Resource availability:** It is the one of the most common benefit facilitated by virtualization. It also helps to track and leverage the resource pool under the same umbrella of resource units.
- **Operational efficiency:** Sometimes conventions and configurations followed by test and operation teams may differ from those followed by development teams. This can cause the application behaviour to be different from what was intended as well as delay services. The template-based approach, with its solution stacks of hardware, configurable applications, and operating systems, is more transparent and can help the teams to understand the environment better.
- **Hosted tools:** Due to these, the developers and testers need not install, configure, run, or maintain tools on their systems as they can log into the tools from any machine on the network maintaining the tools. Rather they can simply login to the tools and enjoy the services over the network.

These four benefits help the developers and testers to concentrate on their core work, retain focus, and concentrate more on their work without worrying about other jobs. This increases quality and productivity, and therefore, more developer innovation, increased test quality and coverage, etc. which are beneficial for an organization.

There are a number of major challenges developers face today in getting started and rolling out new applications and services faster. However, innovative new products and services are the lifeblood of rapidly growing companies. They represent a substantial portion of corporate sales and profits. In an environment of heightened competition, the inability to roll out new applications and services quickly means declining market share and lost revenue.

A growing application backlog leaves lines of business and end-users frustrated because they feel IT is a bottleneck and they look for ways to work around IT to roll out new products and services more quickly. Testing backlog is often very long, and a major factor in the delay of new application deployments.

A major reason testing takes so long is it takes weeks, on average, to set up application environments for test and QA as well as production. This is because of the time it takes to procure new hardware and software, and then schedule time with IT to configure and set up the systems. Configuration and setup are manual processes where errors are easily introduced. The average new application takes six to nine months to deploy, on average. This is caused by a number of factors ranging from poor governance to poor collaboration between business users and development to inflexible infrastructure and tools. Almost 30 percent of all defects are caused by wrongly configured test environments. This is a result of manual processes without any automation to replicate testing environment along with challenges organizations face in finding available resources to perform tests in order to move new applications into production. Test environments are seen as expensive and provide little real business value.

2.4 CLOUD DEPLOYMENT MODELS

Let's talk about cloud computing and the different types of cloud deployment models and different types of services that can be delivered using that model. Cloud computing is a style

of computing in which business processes, application, data, and any type of IT resource can be provided as a service to users.

Cloud delivery models can be briefly classified into three types (Figure 2.2):

- **Public:** In a public cloud, a business rents the capability and they pay for what they use on-demand.
- **Private:** In private clouds, a business essentially turns its IT environment into a cloud and uses it to deliver services to their users.
- **Hybrid:** Hybrid clouds combine elements of public and private clouds.

A private cloud drives efficiency while retaining control and greater customization. Public clouds today are for processes deemed more easily standardized and a lower security risk. There some functions that already exhibit a high degree of standardization, that are more easily

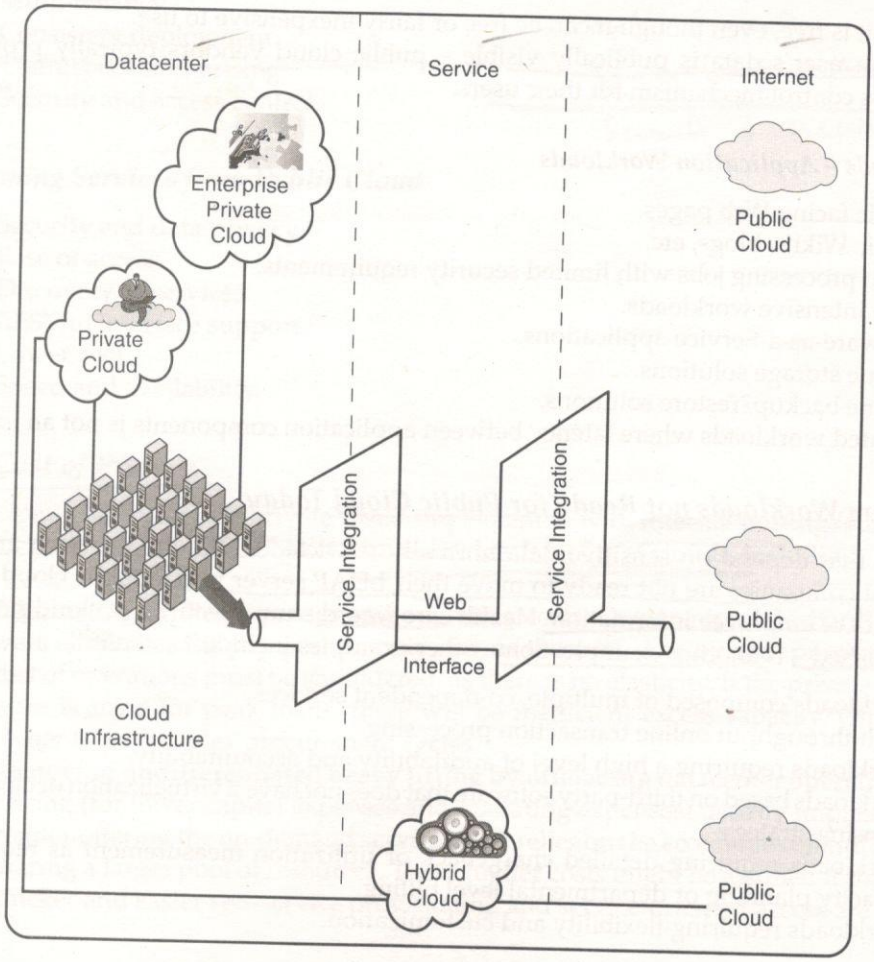

FIGURE 2.2 Private, public, and hybrid clouds.

moved to a public cloud – things such as search, e-commerce, and discreet business processes like sales force management.

There is not a one-size-fits-all model; in a number of cases, businesses may end up using all these models eventually, based on the business model for different services.

2.4.1 Public Clouds

Public cloud services are available to clients from a third-party service provider via the Internet. Public clouds provide an elastic, cost-effective means to deploy solutions and take care of deploying, managing, and securing the infrastructure. Companies can use it on demand, and with the pay-as-you-use option, it is much like utility consumption. Enterprises are able to offload commodity applications to third-party service providers (hosters).

The term 'public' does not mean

- That it is free, even though it can be free or fairly inexpensive to use.
- That a user's data is publically visible – public cloud vendors typically provide an access control mechanism for their users.

Public Clouds – Application Workloads

- Public facing Web pages.
- Public Wiki's, blogs, etc.
- Batch processing jobs with limited security requirements.
- Data intensive workloads.
- Software-as-a-Service applications.
- Online storage solutions.
- Online backup/restore solutions.
- Isolated workloads where latency between application components is not an issue.

Application Workloads not Ready for Public Cloud Today

Workloads that depend on sensitive data normally restricted to the organization are public today. Most companies are not ready to move their LDAP server into a public cloud because of sensitivity of employee information. Health care record – until security of cloud provider is well established – is another example. Some other examples include:

- Workloads composed of multiple, co-dependent services.
- High throughput online transaction processing.
- Workloads requiring a high level of auditability and accountability.
- Workloads based on third-party software that does not have a virtualization or cloud aware licensing strategy.
- Workloads requiring detailed chargeback or utilization measurement as required for capacity planning or departmental level billing.
- Workloads requiring flexibility and customization.

2.4.2 *Private Clouds*

Private clouds are deployments made inside the company's firewall (on-premise datacenters) and traditionally run by on-site servers. Private clouds offer some of the benefits of a public cloud computing environment, such as elastic on-demand capacity, self-service provisioning, and service-based access. They satisfy traditional requirements for greater control of the cloud infrastructure, improving security, and resiliency because user access and the networks used are restricted and designated.

Services in Private Cloud

This section highlights the services provided by private cloud and services consumed by public cloud specifically:

- Virtualization.
- Government and management.
- Multi-tenancy.
- Consistent deployment.
- Chargeback and pricing.
- Security and access control.

Consuming Services from Public Cloud

- Security and data privacy.
- Ease of access.
- Discovery of services.
- RESTful interface support.
- Lower cost.
- Speed and availability.

High 'Cost of Privacy'

Many experts believe that a private cloud implemented with internal hosting/running of the infrastructure makes it difficult to realize many key benefits of clouds, including:

- **Eliminating capital expenses and operating costs:** Ownership of the hardware or software eliminates the pay-per-use potential, as these must be upfront purchases. The full cost of operations must be shouldered, as there is no elasticity. If the private cloud hardware is sized for peak loads, there will be inefficient excess capacity. Otherwise, the owner faces complex procurement cycles.
- **Removing undifferentiated heavy lifting by offloading datacenter operations:** Utility pricing (for lower capital expenses and operating expenses) usually implies an outside vendor offering the on-demand services, and relies on the economies of multiple tenants sharing a larger pool of resources. These higher costs might be justified if the benefits of quicker and easier self-service provisioning and service-oriented access are large.

Private Clouds Provide more Control

In traditional security models, location implies ownership, which, in turn, implies control when security is location-specific. Then location, ownership, and control are aligned. Strong requirements for control and security usually drive a preference for a private cloud, where they own the cloud resources and control the location of those resources. For example, governments may not want their applications or data to reside outside certain borders. Clouds rely on virtualization, and in the public model, this loose coupling breaks the link between location and application, and this reduces the perceived ownership and control.

But control of information is not, in fact, dependent on total ownership or a fixed location. One example is public key encryption – the ownership of the key means control over the information without having to own the rest of the infrastructure. Control can be created over an untrusted infrastructure via a combination of encryption, contracts with service-level agreements, and by (contractually) imposing minimum security standards on the providers. Compliance is difficult outside traditional security models. As long as control through technology and contracts can be clearly demonstrated, it should be possible to make a public cloud computing environment as compliant and as secure as a privately owned facility. Auditors and regulators are continuously adapting to new technologies and business models.

There are two ends to the ownership spectrum – complete implementation ownership, and complete lack of ownership and control of implementation. There are many possible approaches in between, like partial control, shared ownership, etc. There are also different levels of limited access – specific departmental access, industry-only access, controlled partner access, etc.

2.4.3 Hybrid Clouds

A hybrid cloud is a combination of an interoperating public and private cloud. In this model, users typically outsource non-business-critical information and processing to the public cloud, while keeping business-critical services and data in their control. The hybrid model is used by both public and private clouds simultaneously, and is an intermediate step in the evolution process, providing businesses with an on-ramp from their current IT environment into the cloud.

It offers the best of both cloud worlds – the scale and convenience of a public cloud and the control and reliability of on-premises software and infrastructure – and lets them move fluidly between the two based on their needs. This model allows:

- Elasticity, which is the ability to scale capacity up or down in a matter of minutes, without owning the capital expense of the hardware or datacenter.
- Pay-as-you-go pricing.
- Network isolation and secure connectivity as if all the resources were in a privately owned datacenter.
- Gradually move to the public cloud configuration, replicate an entire datacenter, or move anywhere in between.

2.4.4 Community Clouds

A community cloud is controlled and used by a group of organizations that have shared interests, such as specific security requirements or a common mission. The members of the community share access to the data and applications in the cloud.

2.4.5 Shared Private Cloud

This is shared compute capacity with variable usage based pricing to business units that are based on service offerings, accounts datacenters and it requires an internal profit centre to take over or buy infrastructure made available through account consolidations.

2.4.6 Dedicated Private Cloud

Dedicated private cloud has IT Service Catalogue with dynamic provisioning. It depends on Standardized SO architectural assets that can be broadly deployed into new and existing accounts and is a lower cost model.

2.4.7 Dynamic Private Cloud

Dynamic private cloud allows client workloads to dynamically migrate to and from the compute cloud as needed. This model can be shared and dedicated. It delivers on the ultimate value of clouds. This is a very low management model with reliable SLAs and scalability.

2.4.8 Cloud Models Impact

Clouds will transform the IT industry. They will profoundly affect how we live and how businesses operate.

Cloud computing

- Provides massively scalable computing resources from anywhere.
- Simplifies service delivery.
- Provides rapid innovation.
- Provides dynamic platform for next generation datacenters.

Some say it is grids or utility computing or Software-as-a-Service, but it is all of those combined.

Public Clouds: Benefits

There are various ways to benefit from public clouds. Let us see some of the offering facilitated by public clouds:

- Lower barrier to entry/upfront investment.
- Offer self-service for rapid-start development.

- Deliver new pricing models for hardware, software, and service consumption.
- Increase or decrease capacity in minutes.
- Pursue new workloads and opportunities demo/sandbox, collaboration, prototypes.

Internal Private Clouds Drive Cost Savings

There are significant cost savings in implementing an internal private cloud versus a usual traditional infrastructure. With a traditional infrastructure, each server typically runs a single application and the hardware is sized to meet peak demands, which leads to very low average hardware utilization and high software costs due to the number of servers that are deployed and the lack of resource sharing. The internal private cloud uses virtualization on larger servers and leverages advanced service management capabilities to drive efficiency. Servers can be dynamically provisioned to adjust to workload changes and end-users can request the services they need through self-service portals, which drive automation.

Significant cost savings can be achieved by leveraging these capabilities to automate test and development environments. Automation drives down IT labour cost by automatically responding to changes in the environment and taking action before problems occur. Virtualization coupled with service management greatly improves server utilization and reduces software license costs since fewer machines need licenses. Automated provisioning and standardization allows systems to be provisioned in minutes by scripting the install process. In addition, end-users can now interface with IT through self-service portals to request services much like ATMs are leveraged to improve banking service. It can:

- Reduce IT labour cost by 50 percent in configuration, operations, management, and monitoring.
- Improve capital utilization by 75 percent, significantly reducing license costs.
- Lower administrative costs by 50 percent.
- Reduce end-user IT support costs by up to 40 percent.
- Reduce provisioning cycle times from weeks to minutes.
- Benefits of cloud economics with security within your firewall.
- Provide self-service for rapid-start development.
- Provide consistency of application environments.

2.4.9 *Savings and Cost Metrics*

Cloud computing's use of virtualization consolidates systems, which will drive reductions in hardware costs. This is often the initial appeal of funding virtualization projects.

Labour savings are even greater. Many companies still undertake the manual provisioning of IT systems, suffer long and costly delays while people wait for resources to become available, and distract highly skilled personnel from key project to focus on the mundane administration of systems. The automation of these tasks in a highly virtualized cloud environment can save significant labour costs while improving quality and productivity.

The total savings substantially off-set the small incremental increase in software costs that are usually necessary to deliver virtualization and the service management component that are elements of every cloud computing environment.

Cloud computing features two delivery models, private cloud computing and public cloud computing. Private cloud computing exists behind the firewall, while public cloud computing is accessed through the Internet. Cloud vendors believe that these three models – traditional IT, private cloud services, and public cloud services – will all co-exist as part of an overall strategy, based on application type and the business need that would dictate which model.

Hybrid clouds are services delivered to the end user that are composed of both private and public cloud computing elements.

2.4.10 *Commoditization in Cloud Computing*

When businesses started taking advantage of IT, the first organizations to computerize their business processes had significant gains over their competitors. As the IT field matured, the initial competitive benefits of computerization fell. Computerization then became a requirement just to stay on a level playing field. In essence, there is an increasing amount of IT that operates as a commodity.

For example, a paper products company needs a certain amount of unique IT to run its business and make it competitive. But it also runs a huge amount of commodity IT. The commodity technology takes time, money, people, and energy away from their business of producing quality paper products at a competitive price.

As executive management realizes it is operating a lot of commodity IT, which is not core to their competency, the debate shifts from whether cloud computing will take hold in the enterprise to a debate about how much of the organizational IT will be left internal, on-premise. IT functions should be evaluated, and a determination made as to which is a 'commodity' and which is not. Then determine where to place that function in the new IT organization.

2.5 SECURITY IN A PUBLIC CLOUD

Let us now discuss some of the security concerns that should be considered for the cloud deployments.

2.5.1 *Multi-Tenancy*

As long as the cloud provider builds its security to meet the higher-risk client, all of the lower-risk clients get better security than they would have normally. A bandage manufacturer may have a low risk of being a direct target of malfeasants, but a music label that is currently suing file sharers could have a high risk of being targeted by malfeasants. When both the bandage manufacturer and the music label use the same cloud (multi-tenancy), it is possible that attacks directed at the music label could affect the bandage manufacturer's infrastructure as well. So the cloud provider must design the security to meet the needs of the music label – and the bandage manufacture gets the benefits.

2.5.2 Security Assessment

Over time, organizations tend to relax their security posture. To combat a relaxation of security, the cloud provider should perform regular security assessments. The assessments should be done by someone who is experienced and able to identify issues and fix them.

The report should be provided to each client immediately after the assessment is performed so that the clients know the current state of the overall cloud's security.

2.5.3 Shared Risk

Sometimes, a cloud service provider may not be the cloud operator, but may be providing a value-added service on top of another cloud provider's service. For example, if a Software-as-a-Service (SaaS) provider needs infrastructure, it may make more sense to acquire that infrastructure from an Infrastructure-as-a-Service (IaaS) provider rather than building it. These cloud service provider tiers that are built by layering SaaS on top of IaaS, for example, can affect a cloud user's security. In this type of multi-tier service provider arrangement, each party shares the risk of security issues because the risk potentially affects all parties at all layers. This issue must be addressed by taking into consideration the architecture used by the cloud provider and working that information into the total risk mitigation plan.

Prepare strategy to prepare a list

2.5.4 Staff Security Screening

Most organizations employ contractors as part of their workforce. Cloud providers are no exception. As with regular employees, the contractors should go through a full background investigation comparable to the cloud user's own employees.

A cloud provider must be able to provide its policy on background checks and document that all of its employees have had a background check performed as per the policy. The contract between the user and cloud provider should bind the cloud provider to require the same level of due diligence with its contractors.

2.5.5 Distributed Datacenters

Disasters are a fact of life, and include hurricanes, tornadoes, landslides, earthquakes, and even fibre cuts.

In theory, a cloud-computing environment should be less prone to disasters because providers can provide an environment that is geographically distributed. But many organizations sign up for cloud computing services that are not geographically distributed, and therefore, they should require their provider to have a working and regularly tested disaster recovery plan, which includes SLAs.

Organizations that do contract for geographically diverse cloud services should test their cloud provider's ability to respond to a disaster on a regular basis.

2.5.6 Physical Security

Physical external threats should be analyzed carefully when choosing a cloud security provider. Do all of the cloud provider's facilities have the same levels of security? Are you being sold on the most secure facility with no guarantee that your data will actually reside there? Do the facilities have, at a minimum, a mantrap, card or biometric access, surveillance, an onsite guard, and a requirement that all guests be escorted and all non-guarded egress points be equipped with automatic alarms?

2.5.7 Policies

Any organization that says it has never had a security incident is either being deceptive or is unaware of the incidents it has had. It is, therefore, unrealistic to assume a cloud provider will never have an incident. Cloud providers should have incident response policies, and they should have procedures for every client that feed into their overall incident response plan.

2.5.8 Coding

All cloud providers still use in-house software, which may contain application bugs, so every organization should make sure that their cloud provider follows secure coding practices. Also, all codes should be written using a standard methodology that is documented and can be demonstrated to their customer.

2.5.9 Data Leakage

Data leakage has become one of the greatest organizational risks from a security standpoint. Virtually every government worldwide has regulations that mandate protections for certain data types.

The cloud provider should have the ability to map its policy to the security mandate users must comply with and discuss the issues. At a minimum, the data that falls under legislative mandates, or contractual obligation, should be encrypted while in flight and at rest. Further, an yearly risk assessment just on the data in question should be done to make sure the mitigations meet the need. The cloud provider also needs to have a policy that feeds into the security incident policy to deal with any data leakages that might happen.

2.6 PUBLIC VERSUS PRIVATE CLOUDS

A public cloud is a shared cloud computing infrastructure that anyone can access. It provides hardware and virtualization layers that are owned by the vendor and are shared between all customers. It is connected to the public Internet and presents an illusion of infinitely elastic resources.

Initially it does not require upfront capital investment in infrastructure. For consumption-based pricing, the user pays for resources used, allowing for capacity fluctuations over time.

Provisioning is applied through simple Web interface for self-service provisioning of infrastructure capacity. Potentially significant cost savings are possible from providers' economies of scale. Operating costs for the cloud are absorbed in the usage-based pricing. Separate provider has to be found (and paid for) to maintain the computing stack. Users have no say in SLAs or contractual terms and conditions. Sensitive data is shared beyond the corporate firewall. Distance may pose challenges with access performance and user application content for geographic locations. Support for operating system and application stacks may not address the needs of the business.

A private cloud is a cloud computing infrastructure owned by a single party. It provides hardware and virtualization layers that are owned by, or reserved for the business. It, therefore, presents an elastic but finite resource and may or may not be connected to the public Internet.

SLA's contractual terms & condition are negotiable

2.7 CLOUD INFRASTRUCTURE SELF-SERVICE

The cloud infrastructure has to be provisioned and paid for up-front in private clouds. Users pay for resources as used, allowing for capacity fluctuations over time. Self-service provisioning of infrastructure capacity is only possible up to a point in private clouds. Standard capacity planning and purchasing processes are required for major increases. For a large, enterprise-wide solution, some cost savings are possible from providers' economies of scale. The enterprise maintains ongoing operating costs for the cloud, and the cloud vendor may offer a fully managed service (for a price). SLAs and contractual terms and conditions are negotiable between the cloud vendors and customers to meet specific requirements. All data and secure information remains behind the corporate firewall and the option exists for close proximity to non-cloud datacenter resources or to offices if required for performance reasons for geographic locality. Private clouds can be designed for specific operating systems, applications, and use cases, unique to the business.

There is no clear 'right answer', and the choice of cloud model will depend on the application requirements. For example, a public cloud could be ideally suited for development and testing environments, where the ability to provision and decommission capacity at short notice is the primary consideration, while the requirements on SLAs are not particularly strict. Conversely, a private cloud could be more suitable for a production application where the capacity fluctuations are well understood, but security concerns are high.

Cloud computing employs a structured technique to holistically leverage IT industry best practices to uncover areas of relative strength and weakness across multiple IT domains (strategic alignment, computing system and storage, applications and data, processes, organization, finance/environment, and network) to determine readiness for a cloud computing deployment.

Infrastructure strategy and planning for cloud computing strategy gears the clients who are looking for assistance in understanding the business value that the cloud computing model can bring. It is designed to help the clients evaluate their readiness for cloud computing and possible cloud computing uses within their infrastructure. The goal is to develop a high-level vision strategy, value case, and roadmap for cloud computing.

Infrastructure strategy and planning for cloud computing employs a structured technique to holistically leverage IT industry best practices to uncover areas of relative strength and weakness across multiple IT domains to determine readiness for a cloud computing deployment. It is a business and IT executive initiatives to identify where and how cloud computing can drive business value.

2.7.1 *Infrastructure Strategy and Planning Features*

The strategy and planning has three major features:

- Assessment of the current environment to determine strengths, gaps, and readiness.
- Development of the value proposition for cloud computing in the enterprise.
- Strategy, planning, and roadmap to successfully implement the selected cloud delivery model.

Cloud-based systems have brought a new, scalable application delivery service model to the market. Cloud services promise to help reduce capital and operational costs while providing higher service levels. However, cloud services rely heavily on keeping the data and applications they are managing available at all times, and to restore operations quickly following any type of data disaster (database corruption, virus attack, hardware failure,local/regional disaster).

Cloud administrators need to ensure that a minimum of data is at risk by performing backups as frequently as possible, to meet stringent recovery point objectives. And downtime must be limited as well following an outage to meet strict recovery time objectives.

Emerging model where users can have access to applications or compute resources from anywhere with their connected devices through a simplified UI are best suitable alternatives for ease of use. Applications reside in massively scalable datacenters where compute resources can be dynamically provisioned and shared to achieve significant economies of scale. The 'pay-as-you-go' usage model enables users and companies to predict and manage expenses, reduce costs, and simplify operations better.

2.7.2 *The Path to Cloud Computing*

The path from simple virtualization to cloud computing occurs in five somewhat distinct stages.

Stage 1: Server Virtualization

Companies usually start virtualization as a consolidation attempt. The focal point tends to be on reducing capital expenses (like server, storage, and networks), reducing energy costs, and perhaps avoiding or delaying a datacenter build-out or move.

Stage 2: Distributed Virtualization

Once companies start down the virtualization way, and start to achieve capital expense improvements (like server, storage, and networks), the next focus tends to be on elasticity, operational improvements, rapidity, and organizing downtime more efficiently.

Stage 3: Private Cloud

Once processes are designed for alacrity and standards are in place to enable broad automation, the company is ready to look at introducing self-service capabilities based on the virtualization architecture.

Stage 4: Hybrid Cloud

Private clouds will not be the only answer for any enterprise. The self-service portals and interface introduced by private clouds should enable IT enterprises to leverage public cloud services when they make logic without affecting end users.

Stage 5: Public Cloud

Virtualization is not the must thing or are not the stepping stones before companies use public cloud services. Actually, some companies will attempt with cloud in the public cloud arena first, and use their lessons to establish private clouds for their enterprises.

2.8 SUMMARY

We have discussed several models for cloud computing, including private clouds (where the deployment is within the organization's firewall) and public clouds (where the application services and data are hosted by a third party outside the firewall). Consistent data availability and security is a critical success factor for any cloud deployment. Businesses need to ensure that data is adequately protected and can be restored in a timely fashion following any disruption.