[Unit 2: Cloud Service Models]

Cloud Computing (CSC-458)

Nav Raj Vinady

Cloud Service Models:

Cloud Communication:

Cloud communications are Internet-based voice and data communications where telecommunications applications, switching and storage are hosted by a third-party outside of the organization using them, and they are accessed over the public Internet. Cloud services are a broad term, referring primarily to data-center-hosted services that are run and accessed over an Internet infrastructure. Until recently, these services have been data-centric, but with the evolution of VoIP (voice over Internet protocol), voice has become part of the cloud phenomenon.

Cloud communications providers deliver voice & data communications applications and services, hosting them on servers that the providers own and maintain, giving their customers access to the "cloud." Because they only pay for services or applications they use, customers have a more cost-effective, reliable and secure communications environment, without the headaches associated with more conventional PBX system deployment.

Companies can cut costs with cloud communications services without sacrificing features. The success of Google and others as cloud-based providers has demonstrated that a cloud-based platform can be just as effective as a software-based platform, but at a much lower cost. Voice services delivered from the cloud increases the value of hosted telephony, as users can equally well turn to a cloud-based offering instead of relying on a facilities-based service provider for hosted VoIP. This expands their options beyond local or regional carriers.

In the past, businesses have been able to do this for IT services, but not telecom. Cloud communications is attractive because the cloud can now become a platform for voice, data and video. Most hosted services have been built around voice, and are usually referred to as hosted VoIP. The cloud communications environment serves as a platform upon which all these modes can seamlessly work as well as integrate.

There are three trends in enterprise communications pushing users to access the cloud and allowing them to do it from any device they choose, a development traditional IT communications infrastructure was not designed to handle. The first trend is increasingly distributed company operations in branches and home offices, making WANs cumbersome, inefficient and costly. Second, more communications devices need access to enterprise networks – iPhones, printers and VoIP handsets, for example. Third, data centers housing enterprise IT assets and applications are consolidating and are often being located and managed remotely.

Communication-as-a-Service (CaaS):

Communications as a Service (CaaS) is an outsourced enterprise communications solution that can be leased from a single vendor. Such communications can include voice over IP (VoIP or Internet telephony), instant messaging (IM), collaboration and videoconference applications using fixed and mobile devices.

The CaaS vendor is responsible for all hardware and software management and offers guaranteed Quality of Service (QoS). CaaS allows businesses to selectively deploy communications devices and modes on a pay-as-you-go, as-needed basis. This approach eliminates the large capital investment and ongoing overhead for a system whose capacity may often exceed or fall short of current demand.

CaaS service offerings are often bundled and may include integrated access to traditional voice (or VoIP) and data, advanced unified communications functionality such as video calling, web collaboration, chat, realtime presence and unified messaging, a handset, local and long-distance voice services, voice mail, advanced calling features (such as caller ID, threeway and conference calling, etc.) and advanced PBX functionality. A CaaS solution includes redundant switching, network, POP and circuit diversity, customer premises equipment redundancy, and WAN fail-over that specifically addresses the needs of their customers. All VoIP transport components are located in geographically diverse, secure data centers for high availability and survivability.

CaaS offers flexibility and expandability that small and medium-sized business might not otherwise afford, allowing for the addition of devices, modes or coverage on demand. The network capacity and feature set can be changed from day to day if necessary so that functionality keeps pace with demand and resources are not wasted. There is no risk of the system becoming obsolete and requiring periodic major upgrades or replacement.

CaaS requires little to no management oversight from customers. It eliminates the business customer's need for any capital investment in infrastructure, and it eliminates expense for ongoing maintenance and operations overhead for infrastructure. With a CaaS solution, customers are able to leverage enterprise-class communication services without having to build a premises-based solution of their own. This allows those customers to reallocate budget and personnel resources to where their business can best use them.

Advantages of CaaS

From the handset found on each employee's desk to the PC-based software client on employee laptops, to the VoIP private backbone, and all modes in between, every component in a CaaS solution is managed 24/7 by the CaaS vendor. Let's look as some of the advantages of a hosted approach for CaaS;

Hosted and Managed Solutions: Remote management of infrastructure services provided by third parties once seemed an unacceptable situation to most companies. However, over the past decade, with enhanced technology, networking, and software, the attitude has changed. This is, in part, due to cost savings achieved in using those services. However, unlike the "one-off " services offered by specialist providers, CaaS delivers a complete communications solution that is entirely managed by a single vendor. Along with features such as VoIP and unified communications, the integration of core PBX features with advanced functionality is managed by one vendor, who is responsible for all of the integration and delivery of services to users.

<u>Fully Integrated, Enterprise-Class Unified Communications</u>: With CaaS, the vendor provides voice and data access and manages LAN/ WAN, security, routers, email, voice mail, and data storage. By managing the LAN/WAN, the vendor can guarantee consistent quality of service from a user's desktop across the network and back. Advanced unified communications features that are most often a part of a standard CaaS deployment include:

- Chat
- Multimedia conferencing
- Microsoft Outlook integration
- Real-time presence
- "Soft" phones (software-based telephones)
- Video calling
- Unified messaging and mobility

Providers are constantly offering new enhancements (in both performance and features) to their CaaS services. The development process and subsequent introduction of new features in applications is much faster, easier, and more economical than ever before. This is, in large part, because the service provider is doing work that benefits many end users across the provider's scalable platform infrastructure. Because many end users of the provider's service ultimately share this cost (which, from their perspective, is miniscule compared to shouldering the burden alone), services can be offered to individual customers at a cost that is attractive to them.

<u>No Capital Expenses Needed</u>: When business outsource their unified communications needs to a CaaS service provider, the provider supplies a complete solution that fits the company's exact needs. Customers pay a fee (usually billed monthly) for what they use. Customers are not required to purchase equipment, so there is no capital outlay. Bundled in these types of services are ongoing maintenance and upgrade costs, which are incurred by the service provider. The use of CaaS services allows companies the ability to collaborate across any workspace. CaaS can also accelerate decision making within an organization. Innovative unified communications capabilities (such as presence, instant messaging, and rich media services) help ensure that information quickly reaches whoever needs it.

<u>Flexible Capacity and Feature Set</u>: When customers outsource communications services to a CaaS provider, they pay for the features they need when they need them. The service

provider can distribute the cost services and delivery across a large customer base. This makes the use of shared feature functionality more economical for customers to implement. Economies of scale allow service providers enough flexibility that they are not tied to a single vendor investment. They are able to leverage best-of-breed providers such as Avaya, Cisco, Juniper, Microsoft, Nortel and ShoreTel more economically than any independent enterprise

<u>No Risk of Obsolescence</u>: Rapid technology advances, predicted long ago and known as Moore’s law, have brought about product obsolescence in increasingly shorter periods of time. Moore’s law describes a trend he recognized that has held true since the beginning of the use of integrated circuits (ICs) in computing hardware. Since the invention of the integrated circuit in 1958, the number of transistors that can be placed inexpensively on an integrated circuit has increased exponentially, doubling approximately every two years. Unlike IC components, the average life cycles for PBXs and key communications equipment and systems range anywhere from five to 10 years. With the constant introduction of newer models for all sorts of technology (PCs, cell phones, video software and hardware, etc.), these types of products now face much shorter life cycles, sometimes as short as a single year. CaaS vendors must absorb this burden for the user by continuously upgrading the equipment in their offerings to meet changing demands in the marketplace.

<u>No Facilities and Engineering Costs Incurred</u>: CaaS providers host all of the equipment needed to provide their services to their customers, virtually eliminating the need for customers to maintain data center space and facilities. There is no extra expense for the constant power consumption that such a facility would demand. Customers receive the benefit of multiple carrier-grade data centers with full redundancy—and it’s all included in the monthly payment.

<u>Guaranteed Business Continuity</u>: If a catastrophic event occurred at your business’s physical location, would your company disaster recovery plan allow your business to continue operating without a break? If your business experienced a serious or extended communications outage, how long could your company survive? For most businesses, the

answer is “not long.” Distributing risk by using geographically dispersed data centers has become the norm today. It mitigates risk and allows companies in a location hit by a catastrophic event to recover as soon as possible. This process is implemented by CaaS providers because most companies don’t even contemplate voice continuity if catastrophe strikes. Unlike data continuity, eliminating single points of failure for a voice network is usually cost-prohibitive because of the large scale and management complexity of the project. With a CaaS solution, multiple levels of redundancy are built into the system, with no single point of failure.

Infrastructure-as-a-Service (IaaS):

Infrastructure-as-a-Service (IaaS) is the delivery of computer infrastructure (typically a platform virtualization environment) as a service. IaaS leverages significant technology, services, and data center investments to deliver IT as a service to customers. Unlike traditional outsourcing, which requires extensive due diligence, negotiations ad infinitum, and complex, lengthy contract vehicles, IaaS is centered around a model of service delivery that provisions a predefined, standardized infrastructure specifically optimized for the customer’s applications.

Infrastructure as a Service (IaaS) includes the capability to provision processing, storage, networks, and other fundamental computing resources; 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; storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls). Services offered by this paradigm include: server hosting, web servers, storage, computing hardware, operating systems, virtual instances, load balancing, Internet access, and bandwidth provisioning. IaaS clouds often offer additional resources such as; images in a virtual machine image library, raw (block) and file-based storage, firewalls, IP addresses, virtual local area networks (VLANs) and software bundles. Examples of IaaS providers include Amazon CloudFormation, Amazon EC2, Windows Azure Virtual Machines, DynDNS, Google

Compute Engine, HP Cloud, Rackspace Cloud, ReadySpace Cloud Services, and Terremark.

The *IaaS* cloud computing paradigm has a number of characteristics such as: the resources are distributed and support dynamic scaling, it is based on a utility pricing model and variable cost, and the hardware is shared among multiple users. This cloud computing model is particularly useful when the demand is volatile and a new business needs computing resources and it does not want to invest in a computing infrastructure or when an organization is expanding rapidly.

IaaS provides the hardware and the software for servers, storage, networks, including operating systems and storage management software. The Infrastructure as a Service poses the most challenges.

IaaS providers manage the transition and hosting of selected applications on their infrastructure. Provider-owned implementations typically include the following layered components:

- Computer hardware (typically set up as a grid for massive horizontal scalability)
- Computer network (including routers, firewalls, load balancing, etc.)
- Internet connectivity (often on OC 192 backbones)
- Platform virtualization environment for running client-specified virtual machines
- Service-level agreements
- Utility computing billing

Rather than purchasing data center space, servers, software, network equipment, etc., IaaS customers essentially rent those resources as a fully outsourced service. Usually, the service is billed on a monthly basis, just like a utility company bills customers. The customer is charged only for resources consumed. The chief benefits of using this type of outsourced service include:

- Ready access to a preconfigured environment that is generally ITIL-based (The Information Technology Infrastructure Library [ITIL] is a customized framework of best practices designed to promote quality computing services in the IT sector.)
- Use of the latest technology for infrastructure equipment
- Secured, “sand-boxed” (protected and insulated) computing platforms that are usually security monitored for breaches
- Reduced risk by having off-site resources maintained by third parties
- Ability to manage service-demand peaks and valleys
- Lower costs that allow expensing service costs instead of making capital investments
- Reduced time, cost, and complexity in adding new features or capabilities

Some of the most popular IaaS solutions are discussed as below;

<u>Compute as a service</u>: One of the most ubiquitous IaaS offerings today, compute as a service provides compute capacity that includes servers, operating system access, firewalls, routers and load balancing on demand. These systems have management interfaces, and their capacity can be either shared or private. Depending on the provider and the options an enterprise chooses, compute as a service also can include automated patch management, management of infrastructure software, storage management, security management, dedicated customer support and customized SLAs.

<u>Web hosting</u>: Many organizations rely on their websites for marketing and revenue, and any glitch in operations can mean a loss of business. Moving a website to an IaaS based model ensures that the website won’t get bogged down during peak traffic times — and that organizations won’t have to overpay for capacity to manage those traffic spikes. What’s more, loads will always be balanced, and uptime is guaranteed, thanks to SLAs. Other perks include offsite backup and fast connections for eliminating slow page and content downloads, no matter how much rich media a site includes.

<u>Storage as Service</u>: Storage as a Service is a business model in which a large company rents space in their storage infrastructure to a smaller company or individual. In the

enterprise, SaaS vendors are targeting secondary storage applications by promoting SaaS as a convenient way to manage backups. The key advantage to SaaS in the enterprise is in cost savings -- in personnel, in hardware and in physical storage space. For instance, instead of maintaining a large tape library and arranging to vault (store) tapes offsite, a network administrator that used SaaS for backups could specify what data on the network should be backed up and how often it should be backed up. His company would sign a service level agreement (SLA) whereby the SaaS provider agreed to rent storage space on a cost-per-gigabyte-stored and cost-per-data-transfer basis and the company's data would be automatically transferred at the specified time over the storage provider's proprietary wide area network (WAN) or the Internet. If the company's data ever became corrupt or got lost, the network administrator could contact the SaaS provider and request a copy of the data. Storage as a Service is generally seen as a good alternative for a small or mid-sized business that lacks the capital budget and/or technical personnel to implement and maintain their own storage infrastructure. SaaS is also being promoted as a way for all businesses to mitigate risks in disaster recovery, provide long-term retention for records and enhance both business continuity and availability.

<u>Disaster recovery and backup as a service</u>: The idea behind moving disaster recovery to the cloud is to ensure that organizations have uninterrupted access to data and applications, regardless of emergencies, such as power outages, natural disasters or system failures. These solutions always include redundancy and automatic failover to ensure ongoing access, reducing downtime to nearly zero. Many solutions also employ continuous data protection (CDP), which allows for multiple versions of all data sets to be recovered. This gives users the ability to restore data to any point in time. Data and applications are stored in secure offsite facilities. There are two basic options when it comes to disaster recovery as a service: backup and restore from the cloud and backup and restore to the cloud. With the first option, organizations retain applications and data on their own premise, but back up data to the cloud and restore it to hardware on their own premise when a disaster occurs. With the second option, data is restored to virtual machines in the cloud. For mission-critical applications and resources that must be recovered quickly and completely, the best Schoice is often to replicate data to virtual machines.

Desktops as a service: DaaS is, in essence, an IaaS cloud created solely for hosting and serving virtual desktops. Essentially, it's pay-as-you-go computing that allows enterprises to quickly provision, access, run and deactivate virtual desktop machines as needed. Organizations can choose to connect through a private network service instead of the public Internet. In most cases, the service provider offers storage for the virtual computers, ensures security and data protection, and controls the network bandwidth to ensure uptime. Most solutions come with a self-service portal for provisioning and multitenant monitoring, reporting and billing. DaaS is a way to make sure that there are always enough desktop environments available to new workers, with enough storage and all the right applications. And because the desktops can be accessed via the Internet, users can log in and access their familiar workspaces from any location.

Servers as a service: Accessing servers in the cloud means that no matter what the project, or even if it's the busy season, there will always be enough compute power to go around. It's useful for one-time projects that require additional capacity, or for handling spikes in transactions. And because it's a service, enterprises can rest assured that they'll never be paying for more server capacity than they need. Accessing servers as a service also means organizations can cut their IT administrative, maintenance and service workloads. That's particularly important with servers, which can require complex and expensive system administration. The servers are restricted to secure, private areas dedicated to the organization's use, so security is ironclad.

Networking as a service: This is the newest entrant in the IaaS category. The idea is to offer networking resources on demand in order to support virtual networks — resources such as firewalls, load balancing and WAN acceleration services. Simply put, NaaS provides unified connectivity across storage, networking and servers that changes to meet the demands of virtualized infrastructures. In some cases, a networking service can support quality of service (QoS) and other network-based auditing and monitoring services. As with other IaaS services, NaaS involves no upfront costs and supports full scalability, flexibility and security.

Modern On-Demand Computing:

On-demand computing is an increasingly popular enterprise model in which computing resources are made available to the user as needed. Computing resources that are maintained on a user's site are becoming fewer and fewer, while those made available by a service provider are on the rise. The on-demand model evolved to overcome the challenge of being able to meet fluctuating resource demands efficiently. Because demand for computing resources can vary drastically from one time to another, maintaining sufficient resources to meet peak requirements can be costly. Overengineering a solution can be just as adverse as a situation where the enterprise cuts costs by maintaining only minimal computing resources, resulting in insufficient resources to meet peak load requirements. Concepts such as clustered computing, grid computing, utility computing, etc., may all seem very similar to the concept of on-demand computing, but they can be better understood if one thinks of them as building blocks that evolved over time and with techno-evolution to achieve the modern cloud computing model we think of and use today. One example that we can examine is Amazon's Elastic Compute Cloud (Amazon EC2). This is a web service that provides resizable computing capacity in the cloud.

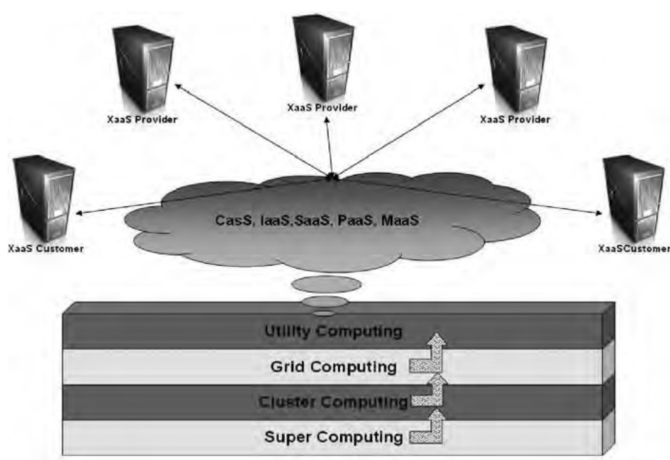

Fig: Building Blocks to the cloud

Amazon Web Services:

Amazon was the first providers of cloud computing (*http://aws.amazon.com*); it announced a limited public beta release of its Elastic Computing platform called *EC2* in August 2006.

AMAZON EC2 SERVICE:

Elastic Compute Cloud (EC2) is a Web service with a simple interface for launching instances of an application under several operating systems, such as several Linux distributions, Microsoft Windows Server 2003 and 2008, OpenSolaris, FreeBSD, and NetBSD.

EC2 allows a user to load instances of an application with a custom application environment, manage networks access permissions, and run the images using as many or as few systems as desired. *EC2* instances boot from an AMI (Amazon Machine Image) digitally signed and stored in *S3*; one could use the few images provided by Amazon or customize an image and store it in *S3*.

A user can,

i. lunch an instance from an existing AMI and terminate an instance;
ii. start and stop an instance;
iii. create a new image;
iv. add tags to identify an image; and
v. reboot an instance.

EC2 is based on the Xen virtualization strategy. In *EC2* each virtual machine functions as a virtual private server and is called an *instance;* an instance specifies the maximum amount of resources available to an application, the interface for that instance, as well as, the cost per hour. This is a web service that provides resizable computing capacity in the cloud. It is designed to make web-scale computing easier for developers and offers many advantages to customers:

- It is a web service interface that allows customers to obtain and configure capacity with minimal effort.
- It provides users with complete control of their (leased) computing resources and lets them run on a proven computing environment.

- It reduces the time required to obtain and boot new server instances to minutes, allowing customers to quickly scale capacity as their computing demands dictate.
- It changes the economics of computing by allowing clients to pay only for capacity they actually use.
- It provides developers the tools needed to build failure-resilient applications and isolate themselves from common failure scenarios.

Amazon EC2 presents a true virtual computing environment, allowing clients to use a web-based interface to obtain and manage services needed to launch one or more instances of a variety of operating systems (OSs). Clients can load the OS environments with their customized applications. They can manage their network's access permissions and run as many or as few systems as needed. In order to use Amazon EC2, clients first need to create an Amazon Machine Image (AMI). This image contains the applications, libraries, data, and associated configuration settings used in the virtual computing environment. Amazon EC2 offers the use of preconfigured images built with templates to get up and running immediately. Once users have defined and configured their AMI, they use the Amazon EC2 tools provided for storing the AMI by uploading the AMI into Amazon S3. Amazon S3 is a repository that provides safe, reliable, and fast access to a client AMI. Before clients can use the AMI, they must use the Amazon EC2 web service to configure security and network access.

A user can interact with *EC2* using a set of SOAP messages, (*The Simple Object Access Protocol (SOAP) is an application protocol developed in 1998 for Web applications. Its message format is based on the Extensible Markup Language. SOAP uses TCP and more recently UDP transport protocols; it can also be stacked above other application layer protocols such as HTTP, SMTP. The processing model of SOAP is based on a network consisting of senders, receivers, intermediaries, message originators, ultimate receivers, and message paths. SOAP is an underlaying layer of Web Services*), and can list available AMI images, boot an instance from an image, terminate an image, display the running instances of a user, display console output, and so on. The user has root access to each instance in the elastic and secure computing environment of *EC2*. The instances can be placed in multiple locations in different Regions and Availability Zones. *EC2* allows the import of virtual machine images from the user environment to an instance through a facility called *VM import.* It also distributes automatically the incoming application traffic among multiple instances using the *elastic load balancing* facility. *EC2* associates an *elastic IP address* with an account; this mechanism allows a user to mask the failure of an instance and re-map a public IP address to any instance of the account, without the need to interact with the software support team.

To be able to connect to a virtual machine in a cloud, a client must know its IP address. For security reasons public IP addresses are mapped internally to private IP addresses. For example, a virtual machine running under Amazon's *EC2* has several IP addresses:

1. *EC2 Private IP Address:* The internal address of an instance; it is only used for routing within the *EC2*Cloud
2. *EC2 Public IP Address:* Network traffic originating outside the *EC2*network must use either the public IP address or the elastic IP address of the instance. The public IP address is translated using the Network Address Translation (NAT) to the private IP address when an instance is launched and it is valid until the instance is terminated. Traffic to the public address is forwarded to the private IP address of the instance.
3. *EC2 Elastic IP Address:* The IP address allocated to an AWS*EC2* account and used by traffic originated outside the *EC2*cloud. NAT is used to map an elastic IP address to the private IP address. Elastic IP addresses allow the cloud user to mask instance or availability zone failures by programmatically re-mapping a public IP addresses to any instance associated with the user's account. This allows fast recovery after a system failure; for example, rather than waiting for a cloud maintenance team to reconfigure or replace the failing host, or waiting for DNS to propagate the new public IP to all of the customers of a Web service hosted by EC2, the Web service provider can re-map the elastic IP address to a replacement instance.

Amazon EC2 Service Characteristics:

- Dynamic Scalability: Amazon EC2 enables users to increase or decrease capacity in a few minutes. Users can invoke a single instance, hundreds of instances, or even thousands of instances simultaneously. Of course, because this is all controlled with web service APIs, an application can automatically scale itself up or down depending on its needs. This type of dynamic scalability is very attractive to enterprise customers because it allows them to meet their customers' demands without having to overbuild their infrastructure.

- Full Control of Instances: Users have complete control of their instances. They have root access to each instance and can interact with them as one would with any machine. Instances can be rebooted remotely using web service APIs. Users also have access to console output of their instances. Once users have set up their account and uploaded their AMI to the Amazon S3 service, they just need to boot that instance. It is possible to start an AMI on any number of instances (or any type) by calling the *RunInstances* API that is provided by Amazon.

- Configuration Flexibility: Configuration settings can vary widely among users. They have the choice of multiple instance types, operating systems, and software packages. Amazon EC2 allows them to select a configuration of memory, CPU, and instance storage that is optimal for their choice of operating system and application. For example, a user's choice of operating systems may also include numerous Linux distributions, Microsoft Windows Server, and even an OpenSolaris environment, all running on virtual servers.

- Integration with Other Amazon Web Services: Amazon EC2 works in conjunction with a variety of other Amazon web services. For example, Amazon Simple Storage Service (Amazon S3), Amazon SimpleDB, Amazon Simple Queue Service (Amazon SQS), and Amazon CloudFront are all integrated to provide a complete solution for computing, query processing, and storage across a wide range of applications. Amazon S3 provides a web services interface that allows users to store and retrieve any amount of data from the Internet at any time, anywhere. It gives developers direct access to the same highly scalable, reliable, fast, inexpensive data storage infrastructure Amazon uses to run its own global network of web sites. The S3 service aims to maximize benefits of scale and to pass those benefits on to developers. Amazon SimpleDB is another web-based service, designed for running queries on structured data stored with the Amazon Simple Storage Service (Amazon S3) in real time. This service works in conjunction with the Amazon Elastic Compute Cloud (Amazon EC2) to provide users the capability to store, process, and query data sets within the cloud environment. These services are designed to make web-scale computing easier and more cost effective for developers. Traditionally, this type of functionality was provided using a clustered relational database that requires a sizable investment. Implementations of this nature brought on more complexity and often required the services of a database administer to maintain it. Amazon Simple Queue Service (Amazon SQS) is a reliable, scalable, hosted queue for storing messages as they pass between computers. Using Amazon SQS, developers can move data between distributed components of applications that perform different tasks without losing messages or requiring 100% availability for each component. Amazon SQS works by exposing Amazon's web-scale messaging infrastructure as a service. Any computer connected to the Internet can add or read messages without the need for having any installed software or special firewall configurations. Components of applications using Amazon SQS can run independently and do not need to be on the same network, developed with the same technologies, or running at the same time. Amazon CloudFront is a web service for content delivery. It integrates with other Amazon web services to distribute content to end users with low latency and high data transfer speeds. Amazon CloudFront delivers content using a global network of edge locations. Requests for objects are automatically routed to the nearest edge server, so content is delivered with the best possible performance. An edge server receives a request from the user's computer and makes a connection to another computer called the origin server, where the application resides. When the origin server fulfills the request, it sends the application's data back to the edge server, which, in turn, forwards the data to the client computer that made the request.

- Reliable and Resilient Performance Amazon Elastic Block Store (EBS): is yet another Amazon EC2 feature that provides users powerful features to build failure-resilient applications. Amazon EBS offers persistent storage for Amazon EC2 instances. Amazon EBS volumes provide “off-instance” storage that persists independently from the life of any instance. Amazon EBS volumes are highly available, highly reliable data shares that can be attached to a running Amazon EC2instance and are exposed to the instance as standard block devices. Amazon EBS volumes are automatically replicated on the back end. The service provides users with the ability to create point-in-time snapshots of their data volumes, which are stored using the Amazon S3 service. These snapshots can be used as a starting point for new Amazon EBS volumes and can protect data indefinitely.

- Support for Use in Geographically Disparate Locations: Amazon EC2 provides users with the ability to place one or more instances in multiple locations. Amazon EC2 locations are composed of Regions (such as North America and Europe) and Availability Zones. Regions consist of one or more Availability Zones, are geographically dispersed, and are in separate geographic areas or countries. Availability Zones are distinct locations that are engineered to be insulated from failures in other Availability Zones and provide inexpensive, low-latency network connectivity toother Availability Zones in the same Region. By launching instances in any one or more of the separate Availability Zones, one can insulate their applications from a single point of failure. Amazon EC2 is currently available in two regions, the United States and Europe.

Elastic IP Addressing:

Elastic IP (EIP) addresses are static IP addresses designed for dynamic cloud computing. An Elastic IP address is associated with your account and not with a particular instance, and you control that address until you choose explicitly to release it. Unlike traditional static IP addresses, however, EIP addresses allow you to mask instance or Availability Zone failures by programmatically remapping your public IP addresses to any instance in your account. Rather than waiting on a technician to reconfigure or replace your host, or waiting for DNS to propagate to all of your customers, Amazon EC2 enables you to work around problems that occur with client instances or client software by quickly remapping their EIP address to another running instance. A significant feature of Elastic IP addressing is that each IP address can be reassigned to a different instance when needed. Now, let’s review how the Elastic IPs work with Amazon EC2 services. First of all, Amazon allows users to allocate up to five Elastic IP addresses per account (which is the default). Each EIP can be assigned to a single instance. When this reassignment occurs, it replaces the normal dynamic IP address used by that instance. By default, each instance starts with a dynamic IP address that is allocated upon startup. Since each instance can have only one external IP address, the instance starts out using the default dynamic IP address. If the EIP in use is assigned to a different instance, a new dynamic IP address is allocated to the vacated address of that instance. Assigning or reassigning an IP to an instance requires only a few minutes. The limitation of designating a single IP at a time is due to the way Network Address Translation (NAT) works. Each instance is mapped to an internal IP address and is also assigned an external (public) address. The public address is mapped to the internal address using Network Address Translation tables (hence, NAT). If two external IP addresses happen to be translated to the same internal IP address, all inbound traffic (in the form of data packets) would arrive without any issues. However, assigning outgoing packets to an external IP address would be very difficult because a determination of which external IP address to use could not be made. This is why implementors have built in the limitation of having only a single external IP address per instance at any one time.

Monitoring-as-a-Service (MaaS):

Monitoring-as-a-Service (MaaS) is the outsourced provisioning of security, primarily on business platforms that leverage the Internet to conduct business. MaaS has become increasingly popular over the last decade. Since the advent of cloud computing, its popularity has, grown even more. Security monitoring involves protecting an enterprise or government client from cyber threats. A security team plays a crucial role in securing and maintaining the confidentiality, integrity, and availability of IT assets. However, time and resource constraints limit security operations and their effectiveness for most companies. This requires constant vigilance over the security infrastructure and critical information assets.

Monitoring as a Service (MaaS) is at present still an emerging piece of the Cloud jigsaw but an integral one for the future. In the same way that businesses realised that their infrastructure and key applications required monitoring tools that would ensure the proactive elimination of any downtime risks, Monitoring as a Service provides the option to offload a large majority of those costs by having it run as a service as opposed to a fully invested in house tool. So for example by logging onto a thin client or central web based dashboard which is hosted by the service provider, the consumer can monitor the status of their key applications regardless of location. Add the advantages of an easy set up and purchasing process and MaaS could be a key pay as you use model for the de-risking of applications that are initially being migrated to the Cloud.

Many industry regulations require organizations to monitor their security environment, server logs, and other information assets to ensure the integrity of these systems. However, conducting effective security monitoring can be a daunting task because it requires advanced technology, skilled security experts, and scalable processes—none of which come cheap. MaaS security monitoring services offer real-time, 24/7 monitoring and nearly immediate incident response across a security infrastructure—they help to protect critical information assets of their customers. Prior to the advent of electronic security systems, security monitoring and response were heavily dependent on human resources and human capabilities, which also limited the accuracy and effectiveness of monitoring efforts. Over the past two decades, the adoption of information technology into facility security systems, and their ability to be connected to security operations centers (SOCs) via corporate networks, has significantly changed that picture. This means two important things: (1) The total cost of ownership (TCO) for traditional SOCs is much higher than for a modern-technology SOC; and (2) achieving lower security operations costs and higher security effectiveness means that modern SOC architecture must use security and IT technology to address security risks.

Protection Against Internal and External Threats

SOC-based security monitoring services can improve the effectiveness of a customer security infrastructure by actively analyzing logs and alerts from infrastructure devices around the clock and in real time. Monitoring teams correlate information from various security devices to provide security analysts with the data they need to eliminate false positives9 and respond to true threats against the enterprise. Having consistent access to the skills needed to maintain the level of service an organization requires for enterprise-level monitoring is a huge issue. The information security team can assess system performance on a periodically recurring basis and provide recommendations for improvements as needed. Typical services provided by many MaaS vendors are described below.

Typical services provided by many MaaS vendors are described below;

<u>Early Detection</u>: An early detection service detects and reports new security vulnerabilities shortly after they appear. Generally, the threats are correlated with thirdparty sources, and an alert or report is issued to customers. This report is usually sent by email to the person designated by the company. Security vulnerability reports, aside from containing a detailed description of the vulnerability and the platforms affected, also include information on the impact the exploitation of this vulnerability would have on the systems or applications previously selected by the company receiving the report. Most often, the report also indicates specific actions to be taken to minimize the effect of the vulnerability, if that is known.

<u>Platform, Control, and Services Monitoring</u>: Platform, control, and services monitoring is often implemented as a dashboard interface and makes it possible to know the operational status of the platform being monitored at any time. It is accessible from a web interface, making remote access possible. Each operational element that is monitored usually provides an operational status indicator, always taking into account the critical impact of each element. This service aids in determining which elements may be operating at or near capacity or beyond the limits of established parameters. By detecting and identifying such problems, preventive measures can be taken to prevent loss of service.

<u>Intelligent Log Centralization and Analysis</u>: Intelligent log centralization and analysis is a monitoring solution based mainly on the correlation and matching of log entries. Such analysis helps to establish a baseline of operational performance and provides an index of security threat. Alarms can be raised in the event an incident moves the established baseline parameters beyond a stipulated threshold. These types of sophisticated tools are used by a team of security experts who are responsible for incident response once such a

threshold has been crossed and the threat has generated an alarm or warning picked up by security analysts monitoring the systems.

Vulnerabilities Detection and Management: Vulnerabilities detection and management enables automated verification and management of the security level of information systems. The service periodically performs a series of automated tests for the purpose of identifying system weaknesses that may be exposed over the Internet, including the possibility of unauthorized access to administrative services, the existence of services that have not been updated, the detection of vulnerabilities such as phishing, etc. The service performs periodic follow-up of tasks performed by security professionals managing information systems security and provides reports that can be used to implement a plan for continuous improvement of the system's security level.

Continuous System Patching/Upgrade and Fortification: Security posture is enhanced with continuous system patching and upgrading of systems and application software. New patches, updates, and service packs for the equipment's operating system are necessary to maintain adequate security levels and support new versions of installed products. Keeping abreast of all the changes to all the software and hardware requires a committed effort to stay informed and to communicate gaps in security that can appear in installed systems and applications.

Intervention, Forensics, and Help Desk Services: Quick intervention when a threat is detected is crucial to mitigating the effects of a threat. This requires security engineers with ample knowledge in the various technologies and with the ability to support applications as well as infrastructures on a 24/7 basis. MaaS platforms routinely provide this service to their customers. When a detected threat is analyzed, it often requires forensic analysis to determine what it is, how much effort it will take to fix the problem, and what effects are likely to be seen. When problems are encountered, the first thing customers tend to do is pick up the phone. Help desk services provide assistance on questions or issues about the operation of running systems. This service includes assistance in writing failure reports, managing operating problems, etc.

The Traditional On-Premises Model:

The traditional approach of building and running on-premises applications has always been complex, expensive, and risky. Building your own solution has never offered any guarantee of success. Each application was designed to meet specific business requirements. Each solution required a specific set of hardware, an operating system, a database, often a middleware package, email and web servers, etc. Once the hardware and software environment was created, a team of developers had to navigate complex programming development platforms to build their applications. Additionally, a team of network, database, and system management experts was needed to keep everything up and running. Inevitably, a business requirement would force the developers to make a change to the application. The changed application then required new test cycles before being distributed. Large companies often needed specialized facilities to house their data centers. Enormous amounts of electricity also were needed to power the servers as well as to keep the systems cool. Finally, all of this required use of fail-over sites to mirror the data center so that information could be replicated in case of a disaster. Old days, old ways—now, let’s fly into the silver lining of todays cloud.

The New Cloud Model : PaaS

PaaS offers a faster, more cost-effective model for application development and delivery. PaaS provides all the infrastructure needed to run applications over the Internet. Such is the case with companies such as Amazon.com, eBay, Google, iTunes, and YouTube. The new cloud model has made it possible to deliver such new capabilities to new markets via the web browsers. PaaS is based on a metering or subscription model, so users pay only for what they use. PaaS offerings include workflow facilities for application design, application development, testing, deployment, and hosting, as well as application services such as virtual offices, team collaboration, database integration, security, scalability, storage, persistence, state management, dashboard instrumentation, etc.

Platform-as-a-Service (PaaS):

Cloud computing has evolved to include platforms for building and running custom web-based applications, a concept known as Platform-as-a- Service. PaaS is an outgrowth of the SaaS application delivery model. The PaaS model makes all of the facilities required to support the complete life cycle of building and delivering web applications and services entirely available from the Internet, all with no software downloads or installation for developers, IT managers, or end users. Unlike the IaaS model, where developers may create a specific operating system instance with homegrown applications running, PaaS developers are concerned only with web based development and generally do not care what operating system is used. PaaS services allow users to focus on innovation rather than complex infrastructure. Organizations can redirect a significant portion of their budgets to creating applications that provide real business value instead of worrying about all the infrastructure issues in a roll-your-own delivery model. The PaaS model is thus driving a new era of mass innovation. Now, developers around the world can access unlimited computing power. Anyone with an Internet connection can build powerful applications and easily deploy them to users globally.

PaaS provides the capability for consumers to have applications deployed without the burden and cost of buying and managing the hardware and software. In other words these are either consumer created or acquired web applications or services that are entirely accessible from the Internet. Usually created with programming languages and tools supported by the service provider these web applications enable the consumer to have control over the deployed applications and in some circumstances the application-hosting environment but without the complexity of the infrastructure i.e. the servers, operating systems or storage. Offering a quick time to market and services that can be provisioned as an integrated solution over the web, PaaS facilitates immediate business requirements such as application design, development and testing at a fraction of the normal cost.

Key Characteristics of PaaS:

Chief characteristics of PaaS include services to develop, test, deploy, host, and manage applications to support the application development life cycle. Web-based user interface creation tools typically provide some level of support to simplify the creation of user interfaces, based either on common standards such as HTML and JavaScript or on other, proprietary technologies. Supporting a multitenant architecture helps to remove developer concerns regarding the use of the application by many concurrent users. PaaS providers often include services for concurrency management, scalability, fail-over and security. Another characteristic is the integration with web services and databases. Support for Simple Object Access Protocol (SOAP) and other interfaces allows PaaS offerings to create combinations of web services (called mashups) as well as having the ability to access databases and reuse services maintained inside private networks. The ability to form and share code with ad-hoc, predefined, or distributed teams greatly enhance the productivity of PaaS offerings. Integrated PaaS offerings provide an opportunity for developers to have much greater insight into the inner workings of their applications and the behavior of their users by implementing dashboard- like tools to view the inner workings based on measurements such as performance, number of concurrent accesses, etc. Some PaaS offerings leverage this instrumentation to enable pay-per-use billing models

Software-as-a-Service:

The traditional model of software distribution, in which software is purchased for and installed on personal computers, is sometimes referred to as Software-as-a-Product. Software-as-a-Service is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet. SaaS is becoming an increasingly prevalent delivery model as underlying technologies that support web services and service-oriented architecture (SOA) mature and new developmental approaches become popular. SaaS is also often associated with a pay-as-you-go subscription licensing model. Meanwhile, broadband service has become increasingly available to support user access from more areas around the world.

International Data Corporation identifies two slightly different delivery models for SaaS. The hosted application management model is similar to an Application Service Provider (ASP) model. Here, an ASP hosts commercially available software for customers and delivers it over the Internet. The other model is a software on demand model where the provider gives customers network-based access to a single copy of an application created specifically for SaaS distribution.

SaaS is most often implemented to provide business software functionality to enterprise customers at a low cost while allowing those customers to obtain the same benefits of commercially licensed, internally operated software without the associated complexity of installation, management, support, licensing, and high initial cost. Most customers have little interest in the how or why of software implementation, deployment, etc., but all have a need to use software in their work. Many types of software are well suited to the SaaS model (e.g., accounting, customer relationship management, email software, human resources, IT security, IT service management, video conferencing, web analytics, web content management). The distinction between SaaS and earlier applications delivered over the Internet is that SaaS solutions were developed specifically to work within a web browser. The architecture of SaaS-based applications is specifically designed to support many concurrent users (multitenancy) at once. This is a big difference from the traditional client/server or application service provider (ASP)- based solutions that cater to a contained audience. SaaS providers, on the other hand, leverage enormous economies of scale in the deployment, management, support, and maintenance of their offerings.

SaaS Implementation Issues:

Many types of software components and applications frameworks may be employed in the development of SaaS applications. Using new technology found in these modern components and application frameworks can drastically reduce the time to market and cost of converting a traditional on-premises product into a SaaS solution. According to Microsoft, SaaS architectures can be classified into one of four maturity levels whose key attributes are ease of configuration, multitenant efficiency, and scalability. Each level is distinguished from the previous one by the addition of one of these three attributes. The levels described by Microsoft are as follows.

XaaS:

Finally XaaS or ‘anything as a service’ is the delivery of IT as a Service through hybrid Cloud computing and is a reference to either one or a combination of Software as a Service (SaaS), Infrastructure as a Service (IaaS), Platform as a Service (PaaS). communications as a service (CaaS) or monitoring as a service (Maas). XaaS is quickly emerging as a term that is being readily recognized as services that were previously separated on either private or public Clouds are becoming transparent and integrated.