Master of Science in Internetworking

Capstone Project Proposal (MINT-709)

On

Moving production apps to AWS cloud and backup strategies

Presented by

Parul Sharma

Student Id :1709434

Under Supervision of

Tai-Chun Lee

ABSTRACT

This project focuses on understanding the trends and advantages of Cloud Computing. Fundamentals and foundational concepts of cloud computing are understood and examined. The project captures the challenges of the traditional data center. It is then how cloud computing is helping to resolve these challenges. A thorough study of cloud architecture is conducted to understand the benefits and principles of cloud computing. The further in-depth analysis provides an understanding and utilization of various cloud computing service models and deployment models. The project also covers understanding and analysis of services and architecture of the biggest cloud vendor Amazon Web Services. A test application is created locally and migrated to the AWS cloud as part of project implementation. This implementation analyzes why more customers are adopting the cloud and how AWS is a leading cloud vendor. Through the project implementation, the AWS cloud architecture and relevance of various services provided by AWS are learned and examined.

ACKNOWLEDGMENT

To begin with, I am extremely thankful to the almighty for giving me immense strength and vision to complete the project work. This was both intriguing and tough in my field of study, but with perseverance from the inside and assistance from others, I am able to share my project work with you.

I express my sincerest gratitude to my project supervisor Mr. Tai-Chun-Lee, who provided me with his valuable input and guidance at each step of the project. Without his countless recommendations, assistance, and mentoring abilities, this endeavor would not have been accomplished. The application stimulating suggestions and cooperation from him provided me with a lot of confidence and a clear direction to take my work toward completion. His profound understanding of the subject inspired me to produce excellent work.

I owe my deep gratitude to Dr. Mike MacGregor and program coordinator Mr. Shahnawaz Mir for encouraging me, providing an opportunity to undertake this project, and also providing the necessary resources to complete this project. Their guidance and help played a great part in selecting this project.

Finally, I would like to thank my family and friends for their encouragement, patience, and unwavering love, which acted as bedrock in this journey. Last but not least I am grateful to the University of Alberta for giving me this opportunity.

Parul Sharma

Contents

List of Figures

CHAPTER 1

INTRODUCTION

1.1 WHAT IS A DATA CENTER?

Data center is a physical location that includes multiple networking components like servers, computers, storage, routers, switches, firewalls, etc., needed to store customer data and host applications. They are high-energy power physical locations that hold a massive amount of data.

"At its simplest, a data center is a physical facility that organizations use to house their critical applications and data. A data center's design is based on a network of computing and storage resources that enable the delivery of shared applications and data. The key components of a data center design include routers, switches, firewalls, storage systems, servers, and application-delivery controllers." [i]

Traditionally, data centers were built in physical locations hosted by customers, where customers store and manage their data. The customers purchase the physical infrastructure required for storing the data in data centers and independently perform all the actions like maintenance, operations, availability, security, and troubleshooting.

Data centers are very critical locations for a customer as they hold the most important data and applications and are not accessible to everyone. The access is restricted to a data center, and it is not available to everyone for use.

Connectivity to the internet is important for any data center to ensure data reaches the end user. There are specific devices at the data center's perimeter connected to the internet through coaxial or fiber optic cables, just like in any other household modem connection. These specific devices are called routers, providing internet connectivity for a data center through very high bandwidth and redundant coaxial or fiber connections.

Since data center holds very valuable information and data, it is highly important to make sure data centers are resilient and scalable. There should not be a single point of failure. Typically, customers have two data centers to store their data. The primary data center is used for all operations, and the secondary data center is used for backup. In case of failure or disaster recovery, the backup data center serves the end user. High availability in data center design enables IT teams and infrastructure to continue functioning even when the primary data center or some part of that data center is not operating. A highly efficient data center offers the highest levels of redundancy and fault tolerance.

Further, since the data center storage critical data and applications, it is very important to maintain the security of the entire infrastructure and the data itself in the data center. Data centers must be secured against external physical threats to their entire infrastructure. A data center facility is kept secure by physical security measures like a secure location, the building's physical access restrictions, and monitoring systems. Data center IT infrastructures demand strict access control included in every data center architecture in addition to the physical security systems installed within a data center (like cameras, locks, etc.).

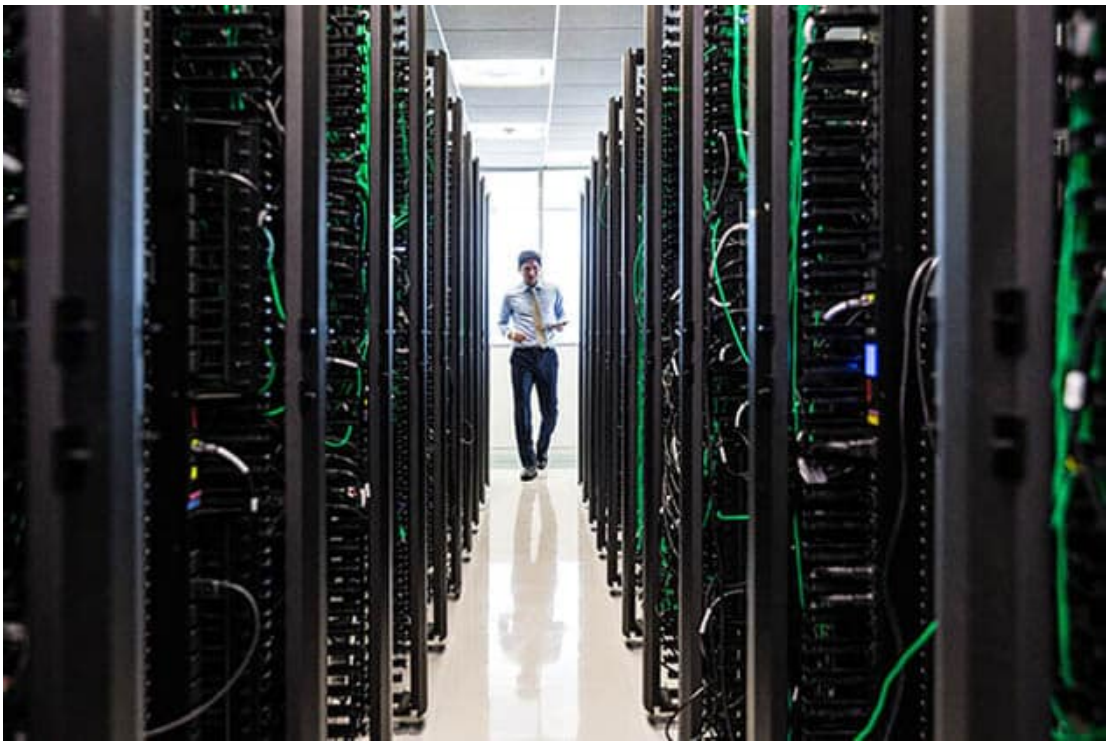

Figure 1 Traditional Data Center [ii]

1.2 CORE COMPONENTS NECESSARY FOR AN EFFICIENT DATA CENTER

Routers, switches, firewalls, storage systems, servers, and application delivery controllers are all part of the design of a data center. Together these components of data centers allow any organization to efficiently process, store and distribute large amounts of data and secure their data. Because these components store and manage business-critical data and applications, the data centers' security is crucial in any data center design.

The core components for any data center include servers, networking, storage, software, cabling infrastructure, power infrastructure, cooling infrastructure, and physical security. All these components work together and are very critical to provide efficient, resilient, and safe data center operations.

These components are clubbed together into below major categories –

- Network Infrastructure – Network Infrastructure is used to provide connectivity between the components and also the external world. Using these components, data is available within an organization and on the Internet. They link storage, data center services, physical and virtualized servers, and external connectivity to end-user locations. Network infrastructure includes routers, switches, and firewalls as the main components. Without network infrastructure, making the data stored in the data center available to the end users is impossible.

- Storage Infrastructure – Data is everything in a data center, and storage components are used to store this data. Data is stored on hard disk drives, tape drives, or any other form of internal and external storage. Storage components are critical for any data center to hold the data and provide backups in case of failure. These components are critical because they hold an organization's most critical and internal information that cannot be leaked or misused.

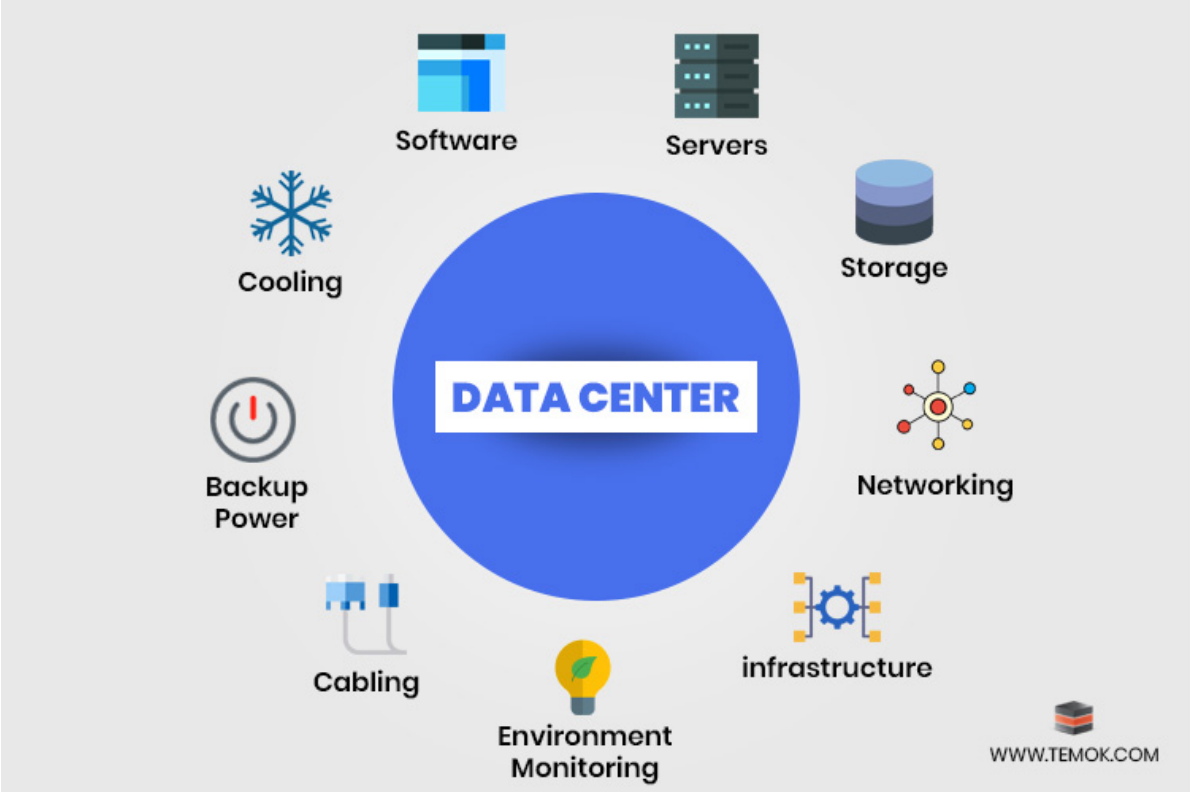

Figure 2 Core components of a Data Center [iii]

- Compute Infrastructure – While data is the fuel of a data center, applications are the engine for the data center to operate. Compute infrastructure provides the necessary computing, memory, processors, local storage, and connectivity to serve the applications.

1.3 WHAT IS CLOUD AND CLOUD COMPUTING?

Cloud, in very simple terms, refers to the “Internet.” Precisely, the cloud refers to whatever you can access remotely via the Internet. When any data is in the cloud, it indicates that it is kept online on the servers on the internet rather than on the hard drives of your computer. [iv]

When you store any data on the internet (“cloud”), it is called data stored on the cloud. The infrastructure required to store the data is available on demand and can be consumed anytime by the user; this is called cloud computing. The infrastructure required to store this data is not directly managed by the data owner. Cloud vendors provide this service in a “pay as you go” model. In simpler terms, cloud vendors maintain and run these data centers and make them available to customers to store their data and applications.

“Pay-as-you-go” is a crucial model in Cloud computing fundamentals, as this is the main payment and consumption method in cloud architecture. In this model, you are only charged for what you use. In other words, you are charged only for the resources you consume, like servers, computing, and storage. This flexible mode of payment makes the cloud a very easy and attractive alternative for customers to store and manage their data. This also makes the cloud a very scalable model for the customers, as they can add resources anytime in the future to meet their requirements and start paying for the resources they are consuming.

Cloud computing refers to the technologies that work together to make the cloud work. The basic underlying technology used in the cloud is “virtualization.” By using special software, IT infrastructure is virtualized, including the —servers, operating system software, networking, and other infrastructure are abstracted so that it can be pooled and divided among the customers, irrespective of physical hardware boundaries. Cloud service companies can utilize their data center resources fully due to virtualization. Without virtualization, cloud computing simply would not exist. These two technologies are intimately intertwined, and virtualization technology is essential to the operation of cloud computing.

Distributing computing services via the Internet (the "cloud"), such as servers, storage, databases, networking, software, analytics, and intelligence, enables quicker innovation, adaptable resource allocation, and scaled economies. Companies can rent anything from application and data storage

to networking infrastructure from a cloud service provider rather than having built their data centers and computing equipment.

One major advantage of employing cloud computing services is that businesses can avoid the upfront costs and complexity of purchasing and maintaining their IT infrastructure by paying for the services they use. On the other hand, by offering the same services to a wide range of consumers, cloud computing service companies can gain enormous economies of scale.

Today, a wide range of services are offered through cloud computing, including storage, networking, processing power, typical office applications, natural language processing, and artificial intelligence.

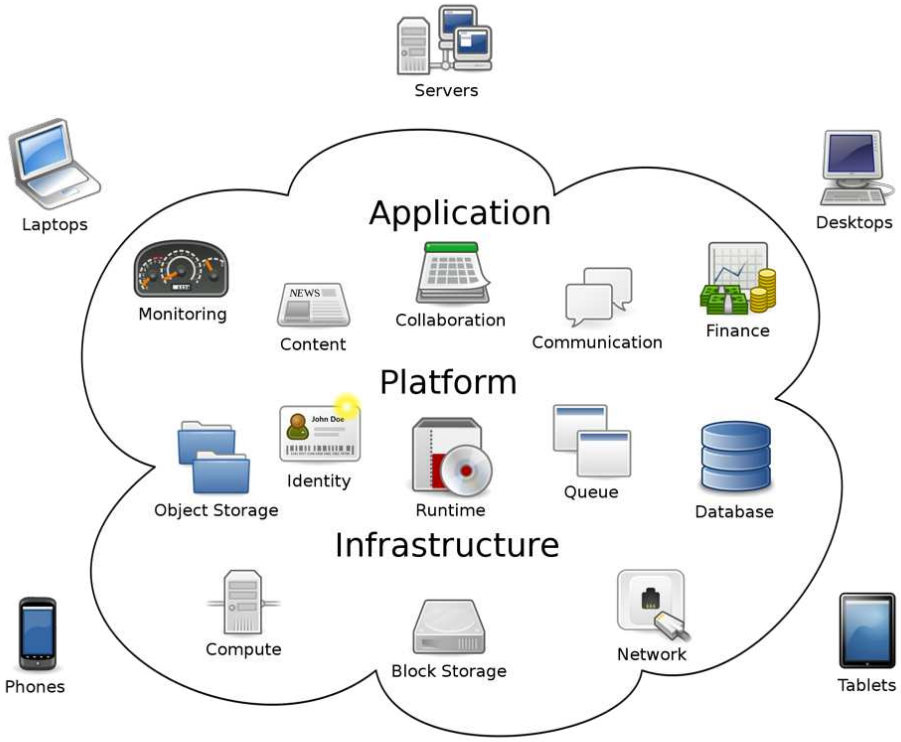

Figure 3 Cloud Computing [v]

Formal definitions of cloud in the industry –

- *Cloud computing is the on-demand availability of computer system resources, especially data storage (cloud storage) and computing power, without direct active management by the user. Large clouds often have functions distributed over multiple locations, each of which is a data center. Cloud computing relies on sharing resources to achieve coherence and typically uses a "pay as you go" model, which can help reduce capital expenses but may also lead to unexpected operating expenses for users.* [vi]

- *Cloud computing is the on-demand delivery of IT resources over the Internet with pay-as-you-go pricing. Instead of buying, owning, and maintaining physical data centers and servers, you can access technology services, such as computing power, storage, and databases, on an as-needed basis from a cloud provider like Amazon Web Services (AWS).*[vii]

- *Simply put, cloud computing delivers computing services—including servers, storage, databases, networking, software, analytics, and intelligence—over the Internet ("the cloud") to offer faster innovation, flexible resources, and economies of scale. You typically pay only for cloud services you use, helping you lower your operating costs, run your infrastructure more efficiently, and scale as your business needs change.* [viii]

1.4 WHY IS IT CALLED CLOUD COMPUTING?

The fundamental reason for using the terms “Cloud” and “Cloud computing” is the location of the data and services. The term "cloud computing" is because the information being accessed can be found remotely on the internet or in another virtual location located elsewhere. In order to convey the idea that the location was unimportant, the analogy of the “cloud” was picked from old telecoms network schematics, in which the public telephone network (which became the internet later on) was often represented as a cloud shape to depict that the location didn't matter – it was just a remote location somewhere out there. [ix] Hence, the term coined was picked for this service-based model to showcase that data is located remotely.

1.5 DIFFERENCE BETWEEN ON-PREMISES DATA CENTER AND A CLOUD DATA CENTER

Essentially, the key difference between a cloud and an on-premises environment is where the data resides. Data in an on-premises environment is present locally, on your business computers and servers, whereas in a cloud environment, data is hosted on the vendor's server and most commonly accessed via a web browser. A collection of servers that are privately owned and managed by companies make up an on-premises data center. On the hand, in cloud computing, businesses lease data center resources from a third-party service provider.

"On-prem" refers to private data centers that companies house in their facilities and maintain themselves. On-prem infrastructure can run private clouds, in which computing resources are virtualized in much the same way as public clouds (however, private clouds can also be run on leased third-party hardware)." [x]

Hence, an on-premises data center refers to a data center that is physically located in an organization's building. The company's IT team keeps the servers up to date and does basic maintenance tasks like installing and updating security software. In contrast, a cloud data center is controlled, managed, and operated by a third-party cloud service provider.

Below are some of the key differences between the operations, ownership, and management of an on-premises data center and a cloud data center-

- In an on-premises environment, the customer maintains their data on their own in private data centers, which are hosted on their campus most of the time. The infrastructure required to maintain that data, which includes the compute, servers, and storage, is all purchased, hosted, and maintained by the customers privately in their environments.

- In an on-premises environment, customers also must always ensure data availability to their users.

- The customer is also responsible for maintaining the redundancy of this data. Typically, customers create a redundant data center, again hosted and maintained by them, to back up all the data.

- A redundant data center can operate in two modes: active-active or active-passive.

- In an active- active mode, both data centers work parallelly, and if one data center fails or is unavailable, the other can be used to continue the operations. In active-passive mode, the primary data center is used for all operations, and the second data center is used as a backup, which comes up only when the primary data center is down.

- In an on-prem environment, the customer is also responsible for maintaining the security of the data on their own. Customers must maintain and configure their authentication and authorization policies and rules to maintain the safety of their data and avoid data leaks and attacks.

- Customers are responsible for day 0, day 1, and day n operations of their data centers hosted in their private environments.

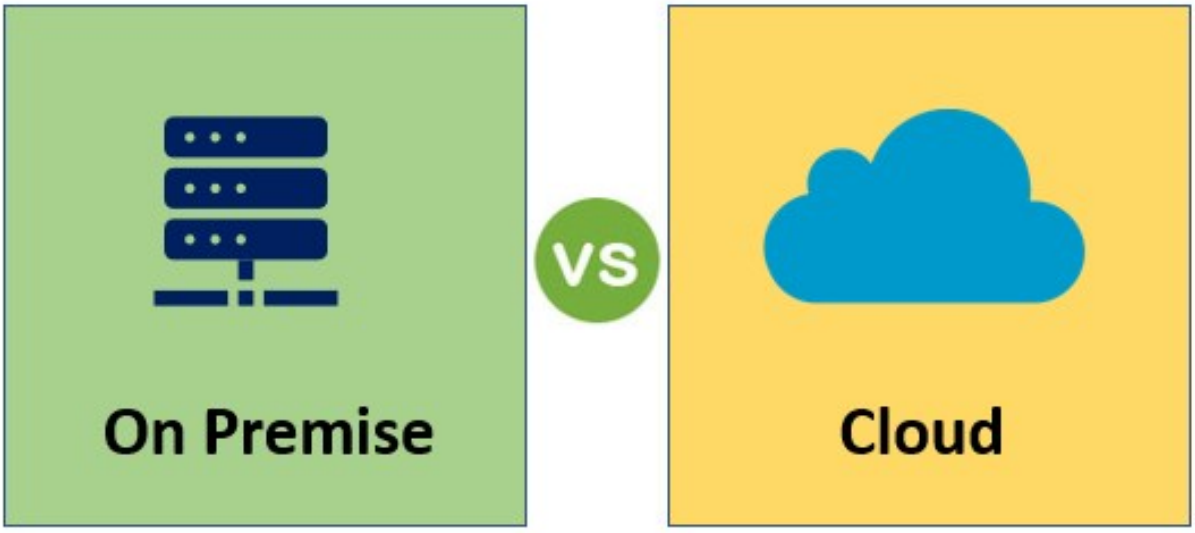

Figure 4 On-prem vs Cloud [xi]

- In contrast to the on-prem environment, the customer does not maintain the cloud environment directly.

- The cloud vendors maintain the cloud environments. Cloud vendors build and run big data centers spread across the globe. They are responsible for maintaining and operating these big data centers.

- Cloud vendors are responsible for day 0, day 1, and day n operations of these big and global data centers.

- Customers don’t have to worry about the underlying infrastructure in a cloud environment, like the compute, storage, and servers, like the maintenance and operations of these data centers.

- These big data centers are then made available to the customers to store and host their resources by the cloud vendors.

- Instead of purchasing infrastructure, the customer just rents these resources from Cloud vendors and consumes them like a service.

- However, these big data centers hosted by cloud vendors are not owned by just one customer. The underlying infrastructure in these big data centers is being shared between multiple customers.

- Although the resources in the cloud environment are shared between multiple customers, their data should be completely segregated from each other and kept private from each other.

- Security in the cloud is generally maintained by cloud vendors and not by the customers. Cloud vendors must follow globally accepted security and compliance rules to maintain the privacy and segmentation of the customer’s data.

- However, the customer is responsible for maintaining their data's authentication and authorization rules.

- Generally, on-prem environments are private data centers that are privately available to the customer.

- On the other hand, cloud environments are called public data centers.

- Essentially the main difference between cloud and on-prem environments is the location of the data. In the cloud, environment data is hosted on the vendor's server and other infrastructure. Whereas in an on-premises environment, data is stored locally in the servers owned by the company.

While considerable differences exist in the operations and management of an on-premises and a cloud data center, they have some similarities. Some of the major similarities between both types of data centers are examined and summarized below –

- Flexible workload management – Both data centers can perform flexible and efficient workload management. Similar to cloud deployments, on-premise data centers can also use technologies and software to streamline workloads. For example, both cloud and on-prem data centers can use machine learning and other testing methodologies to test and run their workloads.

- Access to modern technologies – To ensure security and effectiveness, both on-premises and cloud infrastructure rely on modern technologies. These technologies can include management platforms for observability and troubleshooting, an operating system, and application programming interfaces (APIs), which work together to ensure the smooth operations of the data centers.

- High-level automation – On-premises and cloud environments both use IT infrastructure automation. Automation is a crucial tool for managing large data in both data centers and performing analytics. This facilitates faster workload deployments, cuts costs, and frees up manual and repetitive labor. By implementing automation tools and methods in both types of data centers, customers can save huge amounts of cost in terms of cost and labour hours, removing errors due to manual operations and making significant savings in their budgets.

CHAPTER 2

RESEARCH AND ANALYSIS ON CLOUD ADOPTION AND TRENDS

2.1 EVOLUTION OF CLOUD

Although the name "cloud computing" only came into use in the early 2000s, the idea of "computer as a service" has been around for a very, very long time. The concept of computing can be dated as far back as the 1960s when computer bureaus allowed companies to rent some time on a mainframe for their use rather than have to purchase one themselves. The invention of the personal computer (PC), which made owning the computer much more feasible, largely superseded these "time-sharing" businesses. This was followed by the emergence of corporate data centers, which allowed businesses to store enormous amounts of data. Yet, the idea of renting access to computing capacity has come up repeatedly, most notably in the application service providers, utility computing, and grid computing of the late 1990s and early 2000s. This was followed by the rise of software as a service and hyper-scale cloud-computing companies like Amazon Web Services, after which cloud computing took off. [xii]

DARPA (the Defense Advanced Research Projects Agency) gave MIT $2 million for Project MAC in 1963. The funding for the project contained a stipulation for MIT to develop a technology allowing a “computer to be utilized by two or more people, simultaneously.” In this case, the precursor to what is now frequently referred to as cloud computing was one of those gigantic, outdated machines that used reels of magnetic tape as memory. It functioned as a simple cloud that two or three people could access. This circumstance was referred to as "virtualization," albeit the meaning of the term was later enlarged. [xiii]

The "very" early Internet, also known as the ARPANET (Advanced Research Projects Agency Network), was developed in 1969 with help from J. C. R. Licklider. JCR, also known as "Lick." He was a computer scientist and psychologist who advocated for an idea known as the "Intergalactic Computer Network," in which everyone on Earth would be connected to computers

and have access to information from anywhere. Access to the cloud requires the Intergalactic Computer Network, now known as the internet.

In the 1970s, the meaning of virtualization began to evolve, and now it is used to describe the creation of a virtual machine that performs identically to a real computer. With the advent of the internet, firms started renting out "virtual" private networks, evolving the idea of virtualization. Throughout the 1990s, the usage of virtual computers gained popularity, which prompted the creation of the current cloud computing infrastructure.

Cloud Computing in the Late 1990s

The cloud was initially intended to represent the void between the end user and the provider. According to Professor Ramnath Chellapa of Emory University in 1997, cloud computing is the future "computer paradigm, where the restrictions of computing would be governed by economic rationale, rather than just technical limitations.". This detailed explanation effectively conveys how the cloud formed. The popularity of the cloud increased as companies learned more about its features and advantages. Salesforce emerged as a popular example of a cloud computing application in 1999.

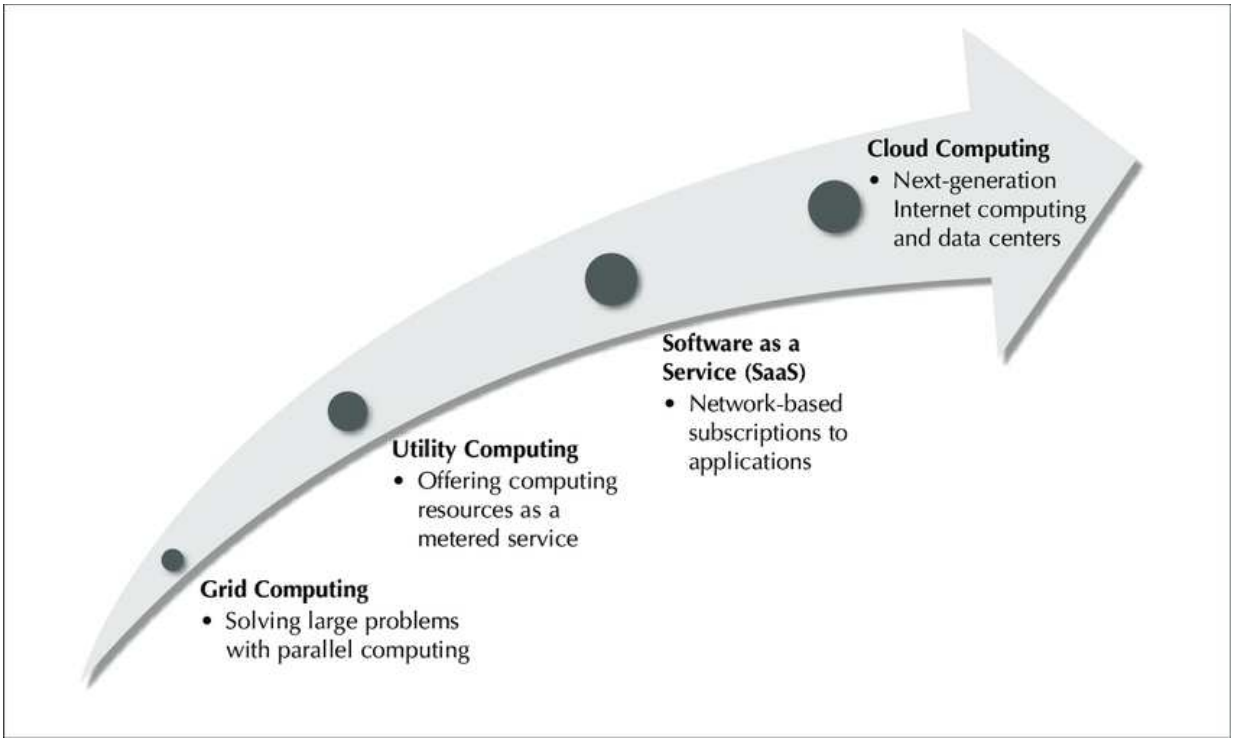

Figure 5 Evolution of Cloud Computing [xiv]

They used it to create the idea of distributing software to users via the Internet. Their program could be accessed and downloaded by anybody with Internet access (or application). Businesses could purchase the software cost-effectively and on-demand without leaving the office.

Cloud Computing in the Early days of the 2000s

In 2002, Amazon (AWS) launched its online shopping services for its customers. It was the first significant company to consider employing only 10% of its capacity, which was typical at the time, an issue that needed to be fixed. They were able to utilize their computer's capabilities far more effectively thanks to the cloud computing infrastructure paradigm. Other major businesses quickly adopted their strategy after seeing AWS's efficiency and success.

Then in 2006, Amazon introduced Amazon Web Services, which provided online services to other websites or customers. Amazon Mechanical Turk, one of Amazon Web Services' websites, offers a range of cloud-based services, including storage, processing, and "human intelligence" to customers. Elastic Compute Cloud (EC2) is another website by Amazon Web Services that enables customers to rent virtual computers and run their software on them.

Google introduced Google Documents services in the same year. Google Spreadsheets and Writely were the foundation for the initial Google Docs version. Writely, which allowed users to save, modify, and import their documents into blogging platforms, was subsequently purchased by Google. A web-based application called Google Spreadsheets on the hand (purchased by Google from 2Web Technologies in 2005) enabled users to create, edit, and share spreadsheets online.

In 2007, IBM, Google, and a number of academic institutions teamed up to create a server farm for research projects that required both quick processors and large amounts of data. The University of Washington was the first organization to sign up and utilize the resources provided by IBM and Google. Following suit shortly after were Carnegie Mellon University, MIT, Stanford University, the University of Maryland, and the University of California, Berkeley. These top universities immediately realized that if IBM and Google supported their research, computer trials might be completed more quickly and for less money. IBM and Google also benefited from the agreement because a lot of the research concentrated on issues they were interested in. Moreover, in 2007, Netflix also introduced its cloud-based streaming video service and offered assistance with "binge-watching."

Further, in 2008, Eucalyptus provided the first Amazon API-compatible platform for delivering private clouds. The first open-source software for setting up private and hybrid clouds was made available that same year by NASA's OpenNebula project. The needs of significant corporations were the focus of many of its most novel aspects.

Cloud Computing in 2010 and Beyond

Private clouds were started in 2008, but they were still underdeveloped and not very popular among customers even in the 2010s. One major factor encouraging the adoption of private clouds was worries about the lack of security in public clouds. Private clouds were established in 2010 by organizations like AWS, Microsoft, and OpenStack and were largely operational. In 2010, OpenStack also made a widely used do-it-yourself cloud open-sourced and free to the general public.

In 2011, the idea of hybrid clouds was first explored. It was necessary to have the capacity to switch workloads between a private and public cloud, as well as some degree of interoperability between the two clouds. Many firms desired to accomplish this then, but relatively few had the systems to do so despite the tools and storage that public clouds could provide. Then IBM released the IBM SmartCloud framework in 2011 to assist its project Smarter Planet. Subsequently, Apple introduced iCloud, which aimed to store more private customer data like photos, music, videos, etc. Further, Microsoft also started to advertise the cloud on television this year to raise awareness about its capacity to store photographs or videos with simple access. The Oracle Cloud, which today offers IaaS (Infrastructure-as-a-Service), PaaS (Platform-as-a-Service), and SaaS (Software-as-a-Service), was released by Oracle back in 2012. Some public clouds provided all of these "basics," while others focused on offering just one. These quickly became the norm. Software as a service gained a lot of traction among big organizations due to its easy and flexibility in consumption.

CloudBolt, founded in 2012, is given the soul credit for developing a hybrid cloud management platform that helped big organizations build, deploy and manage both private and public clouds simultaneously. They fixed the issues with public and private cloud interoperability and brought a new trend in cloud computing.

When businesses began utilizing SaaS providers for specific services like human resources, customer relationship management, and supply chain management, multi-clouds emerged and

became very popular. Around 2013 or 2014, multi-cloud environments gained a lot of popularity. Although using SaaS providers for specific services and benefits is still relatively common, a multi-cloud concept emerged in the market. This idea warned the customers against being pressured to use a specific cloud because of "interoperability difficulties."

By 2014, cloud computing had established its fundamental components; however, security had become a top priority. Because clients valued cloud security so much, it quickly became a very popular service for customers. Cloud security has advanced significantly over the last few years and can now protect par with traditional IT security systems. This covers essential information security from unintentional deletion, theft, and data leakage. Security is, and may always be, the top concern of most cloud users.

App developers are presently one of the major consumers of cloud services. The cloud started to transform in 2016 from being developer-friendly to developer-driven. Application developers started utilizing the cloud's features to their maximum potential. A lot of service work was built with a developer-friendly approach to attract more clients. After realizing the necessity and the potential for profit, cloud suppliers created (and still create) the tools that app developers desire and need.

The Future of Cloud Computing

Cloud has become very famous now. Cloud computing has evolved from its initial model and is much more advanced, feature-rich, and widespread. Almost every organization today uses the cloud for their data or hosts their applications. Cloud has become popular for users to store their data instead of purchasing hard drives or local storage. The coronavirus pandemic has encouraged the growth of online shopping and remote employment. It is reasonable to assume that automated data governance software will be needed in cloud computing to deal with the increasing number of internet laws and regulations.

2.2 WHY COMPANIES ARE ADOPTING CLOUD

Software and services that run remotely over the internet instead of locally on a server or computer at your location are called "data in the cloud." By utilizing the cloud, businesses have found cheaper alternatives while ensuring their clients can access their data and systems from any location. As more computing workloads move to the cloud, whether it be through the use of private clouds made by companies themselves or public clouds offered by suppliers, a sizeable portion of all IT spending is now going towards building the infrastructure to support cloud computing, while spending on conventional, in-house IT is declining. The potential of cloud computing is becoming increasingly apparent to various enterprises, and cloud adoption is growing rapidly and widely. Many cloud providers are available for customers today to migrate from on-prem environments to cloud environments such as AWS, Microsoft Azure, and Google GCP.

A cloud adoption survey conducted by O'Reilly in 2021 also showed that Cloud adoption by companies has increased to 90%. [xv]

Because of this adoption and widespread use of the cloud, it is very important to understand why companies are adopting more and more cloud-based solutions and services. While companies were managing their data previously as well, through its ease of accessibility and cost-effectiveness, the cloud has brought about many new business capabilities which are very attractive to organizations. After researching and understanding these trends, the benefits of cloud computing are broadly categorized into five major sectors, as summarized below –

- Fault Tolerance – One of the most common reasons figured out for organizations to adopt the cloud is the fault tolerance provided by the cloud. In very simple terms, a system's capacity to continue functioning uninterrupted even when one or more of its components fail is known as fault tolerance.

  Any organization that values its continued operation and services always has some disaster recovery measures in its network design. These plans typically include a backup location to store data or your systems if your primary data center has a problem or fails. Traditionally, organizations created a second copy of all their data and applications. This is typically called a secondary data center and is used only for backup in case the primary data center fails. However, the main problem with this approach was that if everything was working with the

primary data center, then the secondary data center was just sitting ideal and was typical of no use. Most of the time, this data becomes obsolete and irrelevant to the company's future business. However, this secondary data center still costs the customer. The customer still pays a huge amount of money to maintain the infrastructure of this secondary data center. This is one of the major sources of expenditure for the customer, which does not bring any value to the customer.

In contrast, cloud architectures have been found to provide customers with much more powerful and flexible backup options by providing backup not in just one server but multiple servers. Customers need only to pay the rent for the server and compute capacity for just storing the dump of their data without spending much money on maintaining the infrastructure like power, electricity, cabling, etc. Customers also have the flexibility to stop these servers anytime the data stored on them becomes obsolete and irrelevant to the company.

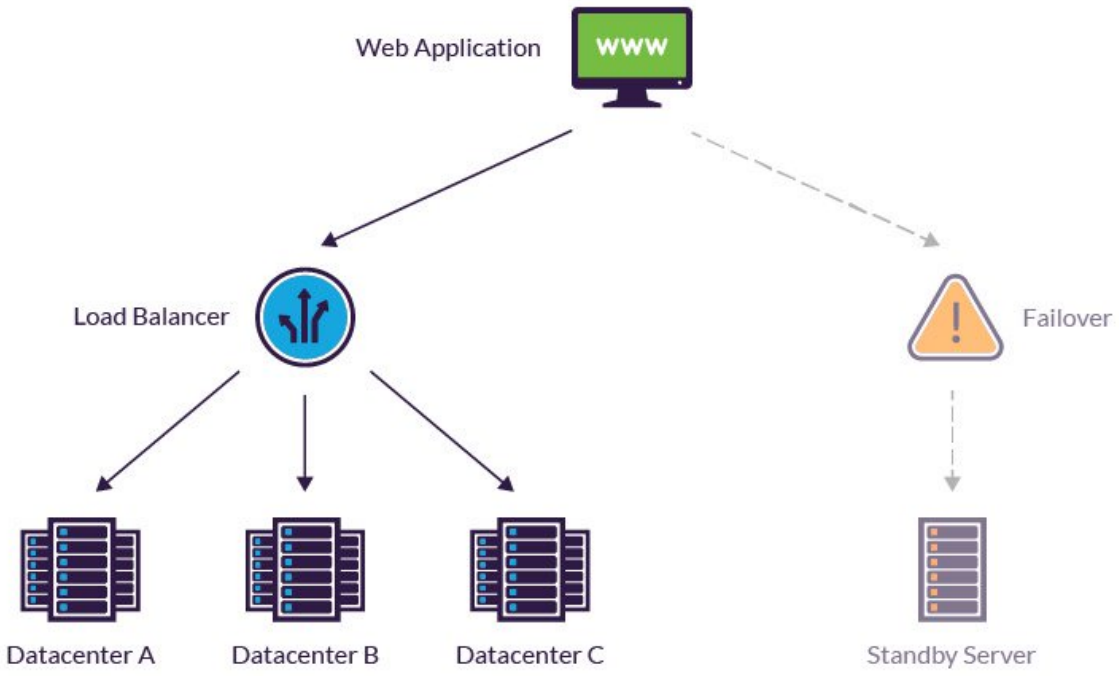

Figure 6 Fault Tolerant architecture [xvi]

- Cost Savings – Research has shown one of the major driven factors for companies to adopt the cloud is significant cost savings. According to one survey, nearly 95% of respondents indicated using the cloud will lower setup and maintenance expenses, and nearly 50% said integrating cloud apps and infrastructure could cut their IT expenditures by up to 50%. [xvii]

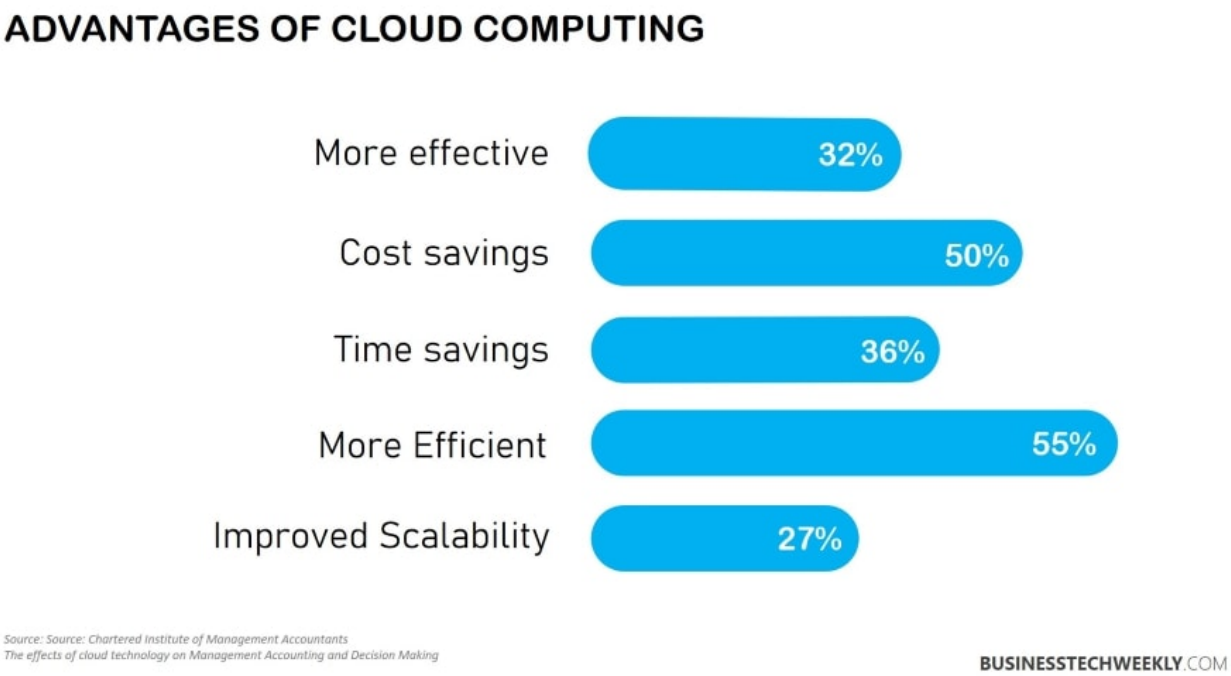

Figure 7 Significant cost savings with cloud computing [xviii]

Cloud computing is helping customers with significant cost savings due to its pay-as-you-go model. With this model, customers pay only for the services they consume. For example, customers only pay for the server capacity they utilize. This is in contrast to a physical data center where customers purchase the entire server even if they are not using the entire capacity of the server. Further, if the requirement for the server or any other services increases in the future, they can add more capacity at any point and start paying the rent according to the updated utilization. This has been recognized as a significant hardware cost-saving mechanism for customers in the cloud.

Further, as customers only rent the resources and do not host them in their environment, the cloud also provides drastic cost savings in infrastructure maintenance. When customers

adopt cloud models, they do not need to pay for infrastructure maintenance costs like power systems, cooling systems, cabling, etc. It has also been observed that with cloud adoption, companies can make significant cost savings while reducing the size of the IT labor needed for maintenance and troubleshooting. This is mainly because the hardware is owned by cloud service providers and kept off-site. As a result, there is less need for internal IT personnel. It does not cost your business any time or money to repair or upgrade servers or other gear; this is the responsibility of the cloud service provider. Additionally, because the cloud provides customers flexibility in terms of usage, customers don't have to plan their capacity and purchase extra resources even if their current requirement is much lower. This further adds to the cost advantage benefits of the customer and has been observed as a major cause of cloud adoption. [xix]

- Scalability – Scalability has been observed as another major reason for cloud adoption. Cloud computing scalability refers to the ability of cloud architecture to scale up or down the capacity of IT resources as required to meet changing and growing business demands. It has been observed that customers wish to maintain their capacity needs as close to their business needs as possible without over-planning or under-planning the capacity. Studies have shown that capacity forecasting is one of the biggest challenges organizations have observed due to predicting the future demands of their business. However, with a cloud environment, customers can add or remove capacity as and when needed to meet their changing business needs. This is possible because the cloud architecture is based on the underlying principle of the "pay-as-you-go" model. Customers have the flexibility and choice to scale up or down based on their business needs, thanks to the pay-as-you-go business model adopted by cloud service providers. Even better, scalability is quick and simple and usually involves little to no downtime. [xx]

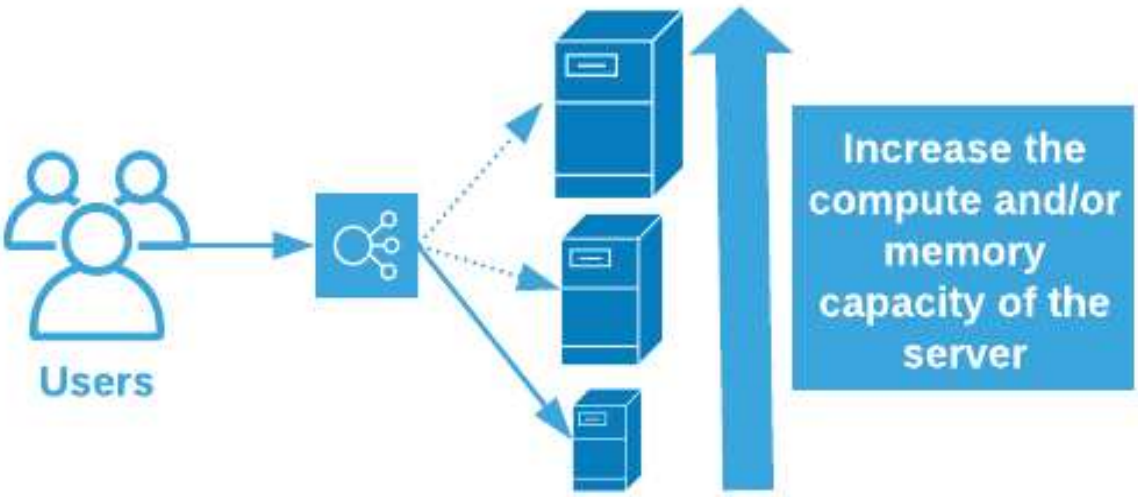

Figure 8 Cloud Computing Scalability [xxi]

- Agility – Agility is a relatively new concept observed in business operation trends. Every company today wants to be more and more agile to drive business growth and quickly adapt to changing business needs. According to the industrial definition, an agile company can react fast to market developments, client and consumer needs, and accelerated timetables. It is capable of making decisions swiftly and acting on them. It can simply transition between projects. [xxii] Because cloud architecture is scalable and flexible in capacity, it can help companies make their IT infrastructure agile according to their business needs. Customers can quickly add more resources and capacity in case the business demand spikes in the future. On the other hand, if the business starts behaving poorly, it can also quickly remove capacity that is no longer being used and save a huge amount of expenditure.

- Globalization - Organizations today have their business operations spread across the world. This trend has further accelerated with the pandemic because users can now work from anywhere in the world. Hence, it has been observed that companies have an increasing demand for a network that is spread across the globe to meet the needs of their business. Most cloud vendors have their network spread all across the globe, and all of their services

can be consumed from any geographic location providing a global reach and seamless experience to the companies. Companies strongly believe this global network is critical for their operation and future growth.

Figure 9 Cloud Architectures are spread across the globe [xxiii]

2.3 WHY WE NEED DIFFERENT CLOUD SERVICES MODELS – SAAS, IAAS, PAAS

Cloud computing has a service-oriented model. In a service-oriented model, consumption and utilization of resources are as-a-service. This concept is popularly known as “Everything is a service.” However, every organization has different needs and requires different services to meet its business goals.

Every company needs different sets of services and different consumption methods to meet its requirements. While one organization would need more control over its consumption model for strict security and compliance reasons, others would prefer a quick and easy service. Each organization has its requirements. These requirements can vary in the amount of control, flexibility, ease of consumption, variations in the feature sets, amount of investment etc. A correct service model is necessary for an organization's business growth. This has been observed as a fundamental reason for building different types of service models for cloud computing.

According to the studies, it was concluded that “One service model cannot fit into the needs” of all businesses. If only one service model is provided to all organizations, it will be ill-equipped for multiple customers because each organization operates on different standards and guiding principles. One model for all might ensure standardization in the industry but will not be able to meet the requirements of the businesses. Rather it is observed all the service models follow core standardization guidelines of cloud computing, like the “pay-as-you-go” model and the concepts of virtualization and globalization.

Hence, the underlying reason for offering different types of service models is to provide differently levels of control, flexibility, and management in each service model to ensure that it meets the company’s business needs. Choosing the right service to meet your business needs is very crucial for the operation and growth of any company. Each model is built to address and meet specific business requirements. Hence, each service model has a different use case and is meant to provide services to a different set of customers. Each model provides customers with a different degree of supervision, visibility, control, flexibility, management, and ease of consumption.

There are three important services models depending on the scope of control, flexibility, and management. Each model has a difference set of relevance and customer base as observed and summarized below. [xxiv]

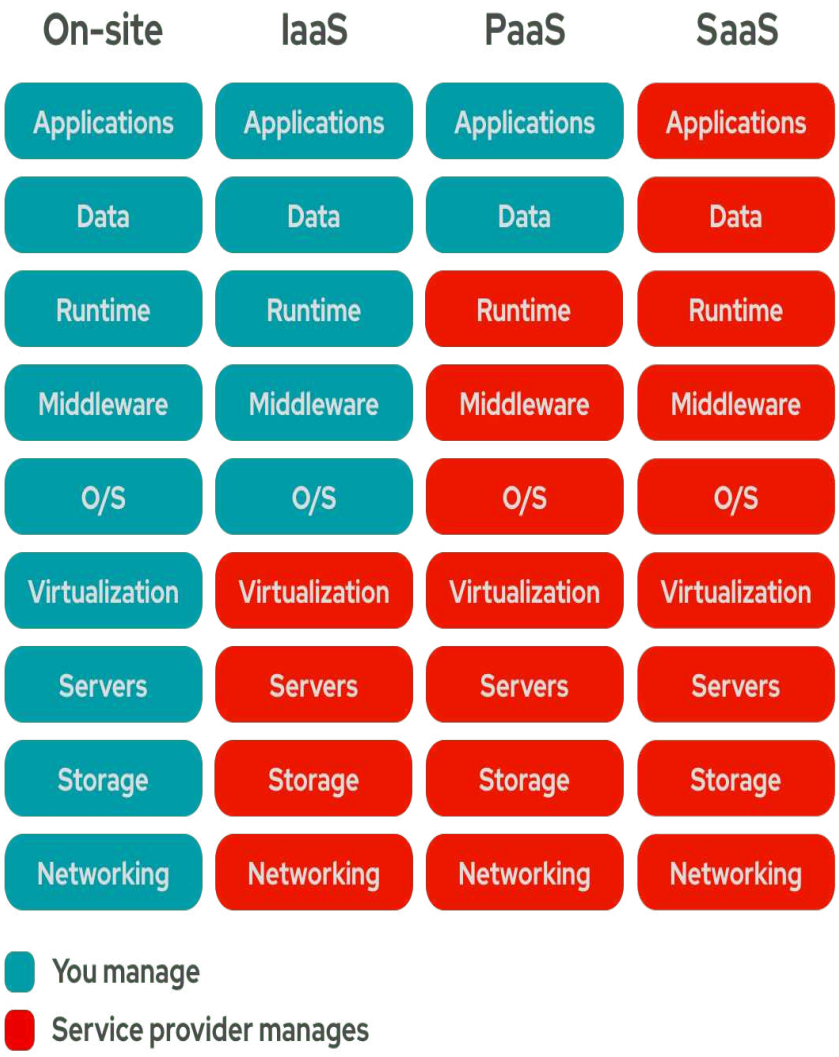

Figure 10 IAAS-vs-PAAS-vs-SAASS [xxv]

- Who should use a Software-as-a-Service Model?
  Software as a service (SaaS) is a cloud-based software solution that allows customers to get an entire application from software providers online. SaaS businesses typically offer access to their products through websites or mobile apps. The SAAS model delivers the entire application to the customer for ready-to-use and direct consumption. In a SAAS

model, the cloud vendor completely manages the application and the underlying infrastructure. The cloud vendor also takes care of the upgrades, maintenance, and support. The vendors take care of the entire computing stack, and the consumers can directly consume the application without worrying about the underlying infrastructure. The most famous example of SAAS software is Gmail and Microsoft 365

Figure 11 SAAS model [xxvi]

SAAS is a ready-to-use service that does not require any installation or deployment. SAAS model is ideal for small or medium size businesses and start-ups who quickly want to consume the service and don't have the necessary budget and resources to spend purchasing the hardware infrastructure and built solutions. [xxvii]

- Who should an Infrastructure-as-a-Service (IAAS) Model?

  The cloud computing service known as "Infrastructure as a Service" (IaaS) provides necessary processing, storage, and networking resources on demand and on a pay-as-you-go basis. Customers can reduce the maintenance of on-premises data centers, save money on hardware, and obtain real-time business insights by moving their organization's infrastructure to an IaaS service. IaaS solutions allow companies to adjust the number of

IT resources according to demand. They also help you provision new apps more quickly and increase the dependability of your underlying infrastructure. IaaS allows you to forego the cost and hassle of acquiring and operating physical servers and datacenter infrastructure. Because each resource is offered as a separate service component, you only pay for a given resource for the time you utilize it. For example, while you buy, install, configure, and manage your software, including operating systems, middleware, and apps, a cloud computing service provider like Azure manages the infrastructure for you and your workloads. [xxviii]

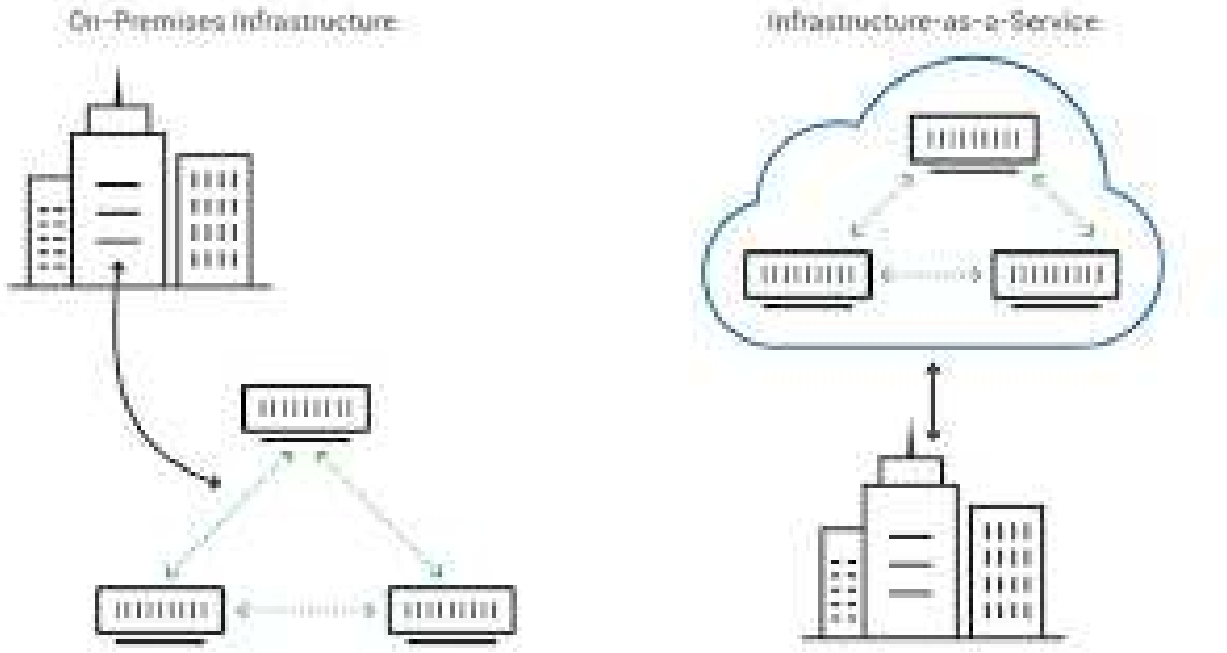

Figure 12 Infrastructure-as-a-Service (IAAS) [xxix]

IAAS is the most flexible cloud services model that allows companies to customize their hardware requirements. IAAS can be used by any type of company, like small businesses, start-ups, or even big enterprises, who just need to rent all essential computing services and have resources, software, and teams available to use this hardware and build solutions. [xxx]

- Who should use Platform-as-a-Service (PAAS) Model?
  Companies that use the Platform-as-a-Service (PaaS) model rent all the resources they need to create an application from a cloud provider. The rent infrastructure, operating systems,

development tools, etc., to build an application. Like IaaS, PaaS also consists of infrastructure, such as servers, storage, and networking, but it also includes other services, such as middleware, development tools, business intelligence (BI) services, database management systems, and other things necessary for application development. The full lifetime of a web application, including development, testing, deployment, management, and upgrading, is supported by PaaS. [xxxi]

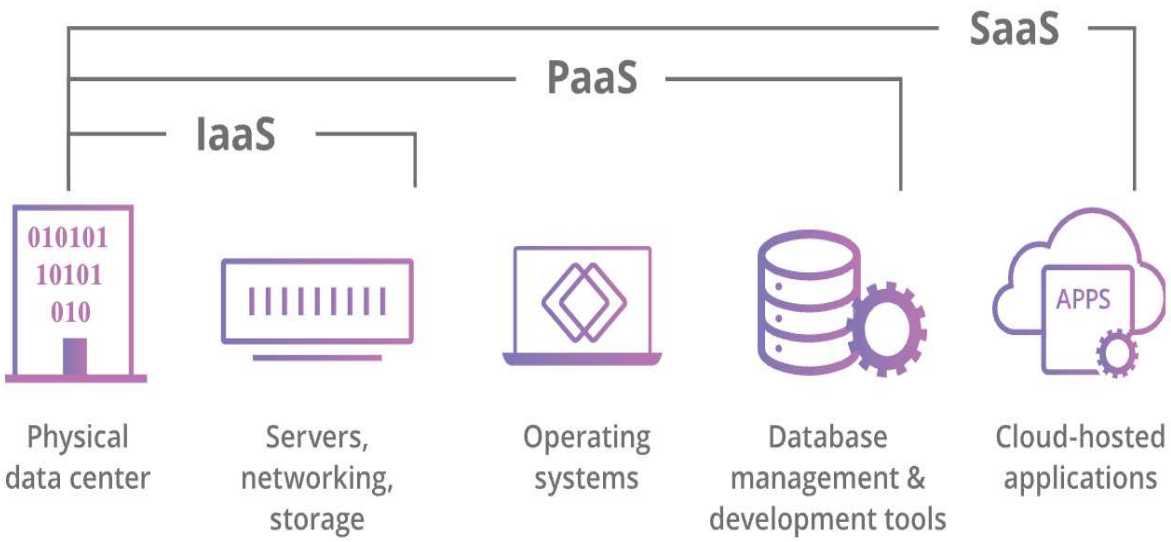

Figure 13 Platform-as-a-service [xxxii]

The most common use case of PAAS is application development. Hence, it is the most suited model for companies that have projects involving developments and multiple vendors. Web Developers are the most common consumers of this type of service model. Developers only need to write their code, while the platform needed to run and implement that code is provided and maintained by the cloud services provider. This can help to significantly reduce the cost and time involved in a development cycle.

According to statistic, the estimated Platform-as-a-Service (PAAS) market is approximately 136 billion USD in 2023. Presently, about 20% of the cloud services market is just by PAAS. [xxxiii]

2.4 WHY WE NEED DIFFERENT CLOUD DEPLOYMENT MODELS – PUBLIC, PRIVATE, HYBRID

After an in-depth study, it was observed that selecting the correct deployment model is very relevant and essential to enjoy the benefits of cloud computing. Choosing the best deployment strategy for the organization is the first stage in the planning process for cloud adoption. The fundamental reason for adopting different cloud deployment models is that every organization has different implementation strategies and requirements, depending on a number of factors such as the number and type of workload and labor available, security needs, budgetary considerations, and the extent of their in-house IT expertise. [xxxiv] In order to meet the implementation needs of different types of organizations, there are broadly three types of deployment models, as explained below –

- Private Cloud - An organization-specific computing infrastructure is known as a private cloud. The data in this deployment model may be kept in the customer's data center or owned by a different service provider. However, the underlying difference in this deployment model is that it is just meant for one specific organization or a customer, and resources are owned and maintained on a private network for that particular customer only. [xxxv]

  Unlike other types of cloud deployment models, the private cloud is not shared between multiple organizations. Security and compliance requirements are the biggest reasons to use private cloud deployment. Customers who use this model generally need a high degree of security and privacy for their data, such as financial organizations and government institutions. They need high control over their data and cannot afford to lose or leak their data. Hence, all their data is completely segregated from other customers and hosted in a private environment, owned and managed by them. These customers might end up renting dedicated cloud services and not sharing them with others. For instance, they can rent an entire server from the cloud vendor instead of just buying a small capacity in that server.

  Advantages of Private cloud – Some of the major advantages observed for customers through this deployment model include a high degree of security, regulatory compliance, higher performance, and granular and more robust control. [xxxvi]

Dis-advantages of Private cloud – Some of the major disadvantages observed for customers with this deployment model are very less flexibility, difficulty in scalability, and high capex and overhead cost in terms of forecasting the resources in advance. [xxxvii]

- Public Cloud – The most common and well-known cloud computing model is the public cloud. The third-party supplier owns and manages all the resources required to run the infrastructure, including servers, storage, networking components, and supporting software. Your users access these resources via the Internet using a web browser. For example, Amazon Web Services, Microsoft Azure, and Google Cloud Platform are the biggest public clouds today. [xxxviii]

  Public clouds are multi-tenant in deployment: where customers share the infrastructure with other organizations, but their data, applications, and workloads are kept completely isolated from others in a safe and secure virtual space to maintain security and compliance.

  Customers can rent the hardware via a subscription-based model and only pay for the services they use, as opposed to owning and maintaining the entire hardware. According to the studies, the public cloud is the most widely accepted deployment model due to its ease of use and exceptional cost savings. Customers are attracted to this model because it provides huge cost savings.

  Advantages of Public cloud - Some of the major advantages observed for customers through this deployment model include the much lower cost in deploying this model, no time and effort required in maintenance, on-demand scalability, high reliability, reduced complexity, and flexible pricing options. [xxxix]

  Dis-advantages of Public cloud – One of the major disadvantages of public cloud deployment is less degree of control, security, and compliance. There is also limited infrastructure visibility and control at times. [xl]

- Hybrid Cloud - In its most basic form, a hybrid cloud blends the two cloud categories, private and public. A hybrid cloud is just a mix of both private and public clouds to get the advantages of both models. In order to enable seamless movement of data and applications between the two types as needs change, high levels of integration and orchestration are necessary for a hybrid cloud deployment model. [xli]

  The primary reason for adopting this model type is to ensure companies can enjoy the best of both worlds. For example, organizations can choose to host their critical data, like details of their clients, in a private cloud, completely owned by them, while keeping their less critical data, like guest information, on the public cloud resources shared with other vendors but still kept isolated from other customers data. More and more companies have been observed adopting the hybrid cloud approach, especially after the pandemic. The major reason is that companies can achieve their technical and business goals more affordably and successfully with a hybrid cloud strategy than with a public or private cloud alone.

  Advantages of Hybrid Cloud – One of the major advantages of hybrid cloud deployment is the huge amount of flexibility customers can get in picking their implementation strategy. It makes them scalable and helps to reduce the cost of operations. [xlii]

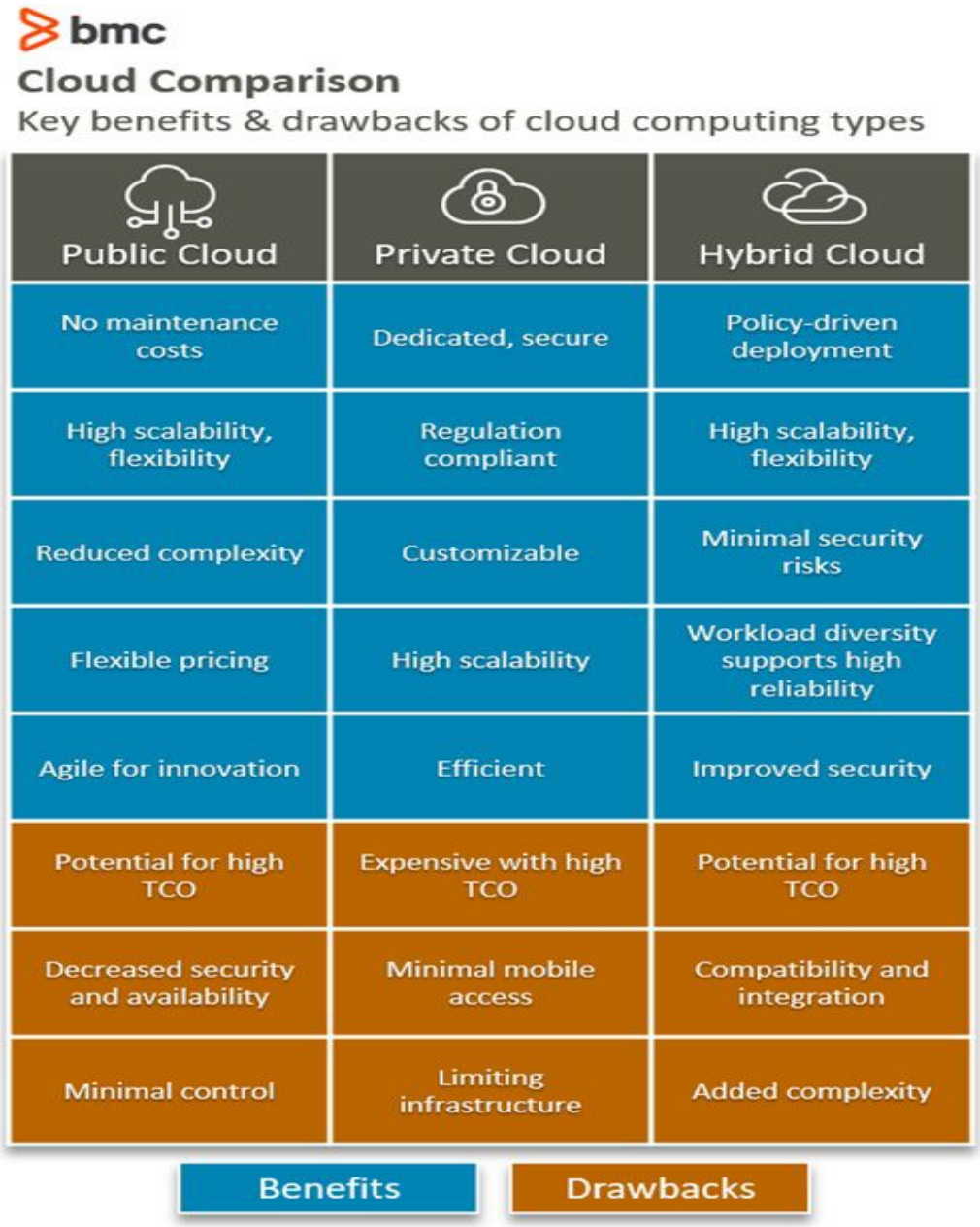

Figure 14 Key Difference between Public, Private and Hybrid Cloud [xliii]

Dis-advantages of Hybrid Cloud – One of the major disadvantages of hybrid cloud deployment is the challenges due to lack of interoperability which can cause difficulty in operations. Another challenge is the lack of strict security and compliance, as the data lies at multiple locations. [xliv]

2.5 RESOLVING THE CONFUSION BETWEEN MULTI- CLOUD AND HYBRID CLOUD

In recent times, the term multi-cloud has become more and more famous. Multi-cloud refers to a model where customers enjoy cloud services from multiple cloud vendors instead of just one cloud vendor. For example, customers can have cloud services from both Amazon and Google Cloud to get the best features from both vendors. [xlv]

Multi-cloud is often confused with a hybrid cloud approach; both are used interchangeably. However, the key difference between both models is the number of cloud vendors and deployment strategy used. Hybrid cloud involves using both public and private cloud deployment models. Multi-cloud, on the other hand, refers to the cloud services consumed by multiple cloud providers. One major challenge in the Multi-cloud approach, like a Hybrid Cloud model, is the difficulty in inter-operability between different cloud vendors.

2.6 BENEFITS OF CLOUD ADOPTION

Cloud computing has become very attractive for customers due to the ease, flexibility and scalability customers can enjoy with cloud architectures in contrast to their traditional model-based on-prem data centers. After thorough analysis, the advantages of cloud adoption are broadly categorized and summarized below –

- Cost Saving – One of the biggest and the most important reasons for customers to adopt the cloud is the significant savings in cost. Cloud follows a pay-as-you-go and subscription-based pricing model, giving customers the ease and flexibility to pay only for the resources they consume. Additionally, customers don't need to pay for the maintenance of the infrastructure. They can directly rent the infrastructure from the cloud vendors and save huge money spent on maintenance, upgrading, troubleshooting, powering, cabling etc.

- Flexibility – Flexibility is the second most important benefit of cloud adoption. Customers adopt more and more clouds because the cloud gives them the flexibility to add and remove resources as per their needs. There is not need for customers to plan their consumption in advance. Additionally, customers can choose from different options available and modify and upgrade as needed to meet their business needs.

- Efficiency – Companies can build, launch, and get their applications to reach markets much quicker and more efficiently without worrying about the underlying infrastructure. This is a great strategic value for customers.

- Scalability – Cloud architectures are completely scalable, allowing customers to add resources quickly to meet their growing business needs. This adds to a huge benefit for customers allowing them to remain as close to their business needs as possible. This also makes businesses more agile, allowing them to adopt to the changes faster and more efficiently.

- Innovation – Cloud computing allows customers to adopt new innovative solutions and capabilities much quicker instead of building in-house resources in their local environments. Also, because customers now do not have the burden of maintaining the infrastructure, they have enough time and resources to drive innovation.

2.7 DISADVANTAGES OF CLOUD ADOPTION

While adopting the cloud provides customers with many benefits and advantages to driving their business, not all things about the cloud are great. Cloud computing comes with some in-built challenges and concerns, so sometimes, customers cannot enjoy the benefits of the cloud models. After thorough analysis, below are the top three dis-advantages of cloud adoption -

- Security – Although cloud computing architecture comes with many inbuilt security capabilities like segmentation, IAM policies, and role-based access controls, since data is lying on the cloud, i.e., on the internet, it is always vulnerable to attacks and thefts. If there is even a single gap in the security policy of the customer using cloud computing, it can cause major security breaches. Further, cloud solutions do not provide the comprehensive and advanced security protections needed today to protect workloads on the internet. These security products must be purchased from other vendors adding to additional costs and gaps.

- Complexity – Another major challenge for a customer in adopting the cloud is the complexity of the architecture. While multiple choices and options are available to the customers, most of the time, customers are unaware of the best possible solutions for their environment and end up selecting products and services which add additional complexity and create vulnerability gaps in the architecture.

- Migration and interoperability – The third major challenge observed with cloud computing is the complex migration cycles and difficulty in interoperability with on-prem resources. In the earlier days, it was also difficult to interoperate between different cloud vendors. However, with increasingly adopting multi-cloud environments, these challenges are being resolved to some extent. However, some vendors still provide solutions that are incompatible with those of other cloud vendors, creating a lack of flexibility and interoperability challenges for customers.

2.8 MARKET ANALYSIS ON TOP CLOUD VENDORS IN THE MARKET TODAY

There are multiple cloud providers in the market today, considering the demand for cloud by the organization and network transformation. However, research has shown below that the top 10 cloud providers provide most of the cloud services to customers. Most of these vendors have their reach globally and provide cloud services worldwide. The difference lies in their scale, ease of availability, the vastness of the services, flexibility, pricing, operating standards, security polices etc. However, all the cloud vendors follow the "As-a-Service" model to provide cloud features and functionalities to the customers. [xlvi]

1. Amazon Web services
2. Google cloud platform
3. Microsoft Azure
4. IBM Cloud
5. Oracle cloud
6. Alibaba cloud
7. Salesforce
8. Verizon Cloud
9. VMware
10. RedHat

Figure 15 Top Cloud Vendors [xlvii]

Amazon Web Services (AWS) leads the world in terms of cloud services. By revenue and market share, Microsoft Azure and Google Cloud Platform are the two largest cloud providers after AWS. These three companies account for over 65% of cloud services provided today to customers. [xlviii]

Cloud Service Provider	Market Share
Amazon Web Services	34%
Microsoft Azure	22%
Google Cloud Platform	9.5%
Alibaba Cloud	6%
Oracle Cloud	2%

Figure 16 Cloud Vendors Market Share [xlix]

CHAPTER 3

AMAZON WEB SERVICES (AWS):

INFRASTRUCTURE AND SERVICES ANALYSIS

3.1 WHAT ARE AMAZON WEB SERVICES AND AWS CLOUD

Amazon Web Services, commonly called AWS, is one of the world's leading cloud vendors. The parent company for AWS is Amazon (amazon.com), which provides many other services like amazon music, amazon pay, amazon prime, amazon Alexa etc. [1]AWS is the world's biggest cloud vendor and has data centers spread across the globe. The most comprehensive and widely used cloud platform in the world, Amazon Web Services (AWS), provides more than 200 fully functional services from its large fleet of data centers worldwide. [li] [lii]

AWS is one of the world's most prominent organizations. It started its web services operation in 2002 (almost 20 years ago) and cloud computing services in 2006 (nearly 16 years). AWS has transformed how networks operate by providing extremely powerful cloud services.

Figure 17 Amazon Web Services [liii]

AMAZON WEB SERVICES CLOUD, commonly referred to as AWS CLOUD, is a collection of large numbers of extremely powerful data centers built and maintained by AWS. AWS provides a vast number of services to customers running on these data centers through the internet. These services are available on-demand with pay-as-go-pricing. [liv]

The AWS Cloud offers a wide range of cloud-based products in the following categories: computing, storage, databases, analytics, networking, mobile, developer tools, management tools, the Internet of Things, security, and corporate applications, which are available on-demand and accessible in a matter of seconds.

Some of the world's most famous services, like Netflix, Twitter, WhatsApp, and many other major organizations, run on AWS cloud today. These organizations rely on AWS for their entire infrastructure, from computing, servers, storage, security, networking, etc.

3.2 NOTABLE CUSTOMER BASE OF AWS

Currently, AWS is the world's leading cloud services provider, and they have their data centers spread all across the world. More than 150,000 developers have registered to utilize Amazon Web Services since its launch, according to a press release issued by Amazon on March 14, 2006. [lv]

Because of its global scale and reach, AWS provides cloud services to millions of customers every day. There are more than 1 million users actively using Amazon Web Services. [lvi]

Around 10% of AWS users are enterprise-scale customers, with the remaining 90% being small and medium-sized businesses, according to several consultancy organizations. [lvii]

- NASA, the 2012 Obama presidential campaign, and Netflix are notable clients of AWS. [lviii]
- According to research from 2019, more than 80% of the DAX-listed companies in Germany use AWS. [lix]
- It was revealed in October 2021 that several UK government agencies and intelligence organizations, including GCHQ, MI5, MI6, and the Ministry of Defence, have hired AWS to hold their sensitive data. [lx]

Top Companies that use AWS cloud services-

Company Name	Website	Location	Revenue
Netflix	www.netflix.com	United States	$29.7 Billion
Facebook	www.facebook.com	United States	$119 Billion
LinkedIn	www.linkedin.com	United States	$10 Billion
Twith	www.twitch.tv	United States	$414 Million
Bitdefender	www.bitdefender.com	United States	$316 Million

Figure 18 Companies that use AWS [lxi]

3.3 EVOLUTION OF AMAZON WEB SERVICES

Amazon started its operations in 1994. The founder of the company Jeff Bezos launched Amazon on July 5, 1994, and he specifically chose the Seattle area for his company because of the region's wealth of technical expertise and because Microsoft was operating in that region. [lxii]

Amazon went public in May 1997, and till 1998 Amazon was mainly selling music and videos. [lxiii] Amazon introduced its subsidiary Amazon Web Services (AWS) only in 2002, with the goal of giving web developers access to APIs to create web apps on top of Amazon's e-commerce platform. [lxiv]

There are many myths and stories about how AWS has evolved. However, according to the most accepted story, AWS began as a need of Amazon to launch an e-commerce website called Merchant.com, which could provide different vendors like Target and Marks & Spencer to sell their products through the same e-commerce platform. [lxv]It was a much more difficult task than initially assumed by Amazon, and eventually, to overcome the challenges, amazon started using API to achieve this goal. During this process, the idea of providing infrastructure services and computing as a service for renting came out, and eventually, Amazon Web Services was launched.

While in 2002, AWS was just providing API services, in 2004, AWS was expanded to provide more services like website popularity statistics and web crawler data from the Alexa Web Information Service. In 2006 AWS started providing Simple Storage Services (S3) to its customers, and in 2008 Elastic Compute Cloud was introduced, which allowed companies to rent data storage and computing power from Amazon.

Since then, till today, Amazon Web Service has grown tremendously in its scale, operations, number of services provided, and customer base. [lxvi] In April 2021, AWS reported 32% yearly growth, accounting for 32% of the $41.8 billion cloud market in Q1 2021. [lxvii]

3.4 WHAT IS THE RELEVANCE OF AWS GLOBAL INFRASTRUCTURE FOR CUSTOMER

AWS cloud is the most extensive and widespread cloud in the world, with its data centers spread across the globe. It's the world's most secure, connected, extensive, and reliable cloud platform. AWS provides the world's most extensive global footprint.

AWS cloud is present in - [lxviii]

i. 31 Launched Regions
ii. 99 Availability Zones
iii. 400+ Edge Locations and 13 Regional Edge Caches

As stated by AWS, The AWS Cloud spans 99 Availability Zones within 31 geographic regions around the world, with announced plans for 12 more Availability Zones and 4 more AWS Regions in Canada, Israel, New Zealand, and Thailand. [lxix]

After an in-depth analysis of the architecture, it was observed that the global infrastructure of AWS is very important for the customers, as it gives them the capability and flexibility on how and in which part of the world they want to run their workloads and applications. This global reach is very relevant for companies with a global reach in their business. If any organization wants to run its services globally, it can pick from any global services provided by AWS. On the other hand, if any company has its operations in one particular location only, it can also pick and choose from the regional and local services provided by AWS.

In addition, to the global scale, this world-wide infrastructure of AWS also provides the customer with High-Availability, Resiliency, Load Sharing, and Disaster Recovery capabilities. Customers can back up their data in a data center in the same location or a separate geographic location to ensure complete resiliency in case of disaster or failure of a data center in one location.

Another important use case of this global scale that was figured out is that it allows companies to provide a fast and efficient performance to end users accessing the company's applications or data over the internet. Because all these data centers are present all across the globe and also have caching capabilities, users can access the data much faster and enjoy a similar experience even if they are sitting in a different part of the world.

AWS maintains this global reach and scale through a three-layered global infrastructure as summarized below -

- AWS Region - An AWS region is an actual physical location worldwide. Within these AWS regions, AWS has built clusters of data centers. These clusters of data centers are called Availability Zone (AZ). Each AWS Region is made up of a minimum of three geographically isolated and physically distinct AZs to ensure high availability and uptime all the time. [lxx]

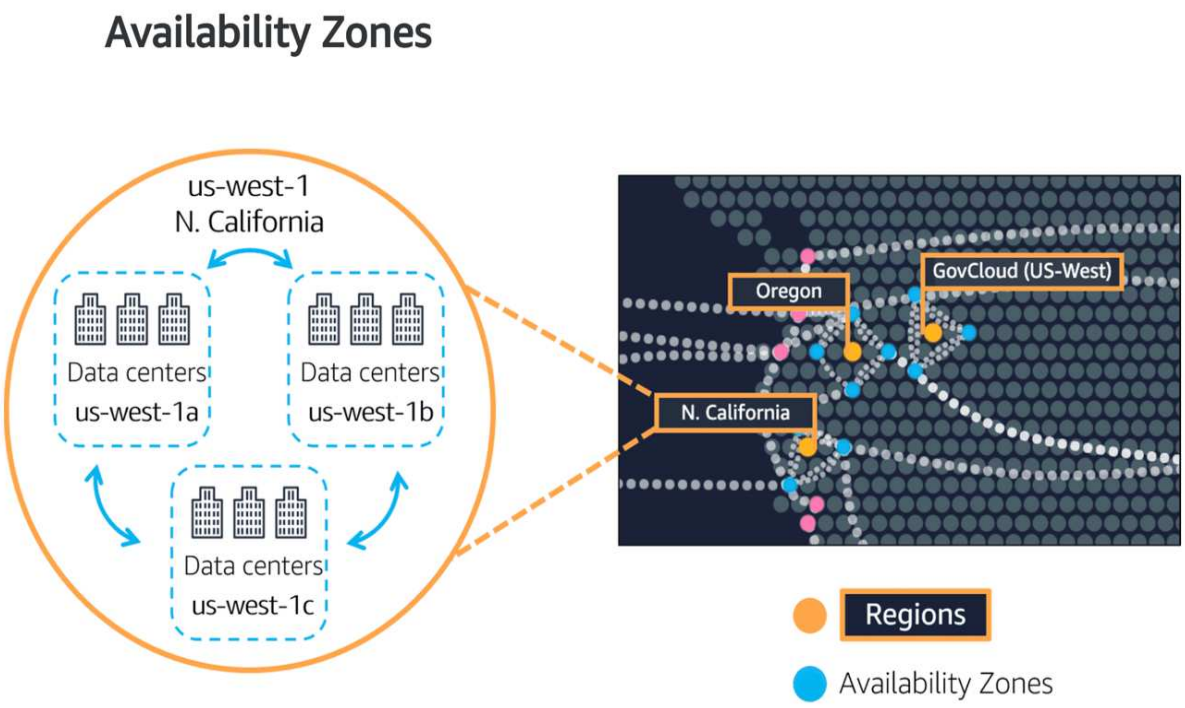

Figure 19 AWS Region and AZ [lxxi]

- AWS Availability Zone - One or more distinct data centers with redundant power, networking, and connectivity make up an Availability Zone (AZ) in an Amazon Region. All the AZs in a region are fully connected over a high-bandwidth, fully redundant, and dedicated microfiber. The data centers with an AZ are also fully connected. They are located close enough to provide a customer with a high availability and low latency network. However, they are far enough from each to avoid the impact of the disaster on

one data center to another. Completely encrypted traffic flows between the AZs, adding an additional layer of security. AZs provide high availability, isolation, and redundancy for customers' data. Businesses are better insulated and shielded from problems like power outages, lightning strikes, tornadoes, earthquakes, and more if an application is partitioned between AZs. Many kilometers physically separate each AZ from every other AZ, even though they are all 100 kilometers (60 miles) apart. [lxxii]

- AWS Edge Location - AWS Edge location is basically a site that AWS uses to cache data closer to the customer. This provides faster delivery of data to the customer. It is a site used by Amazon CloudFront.

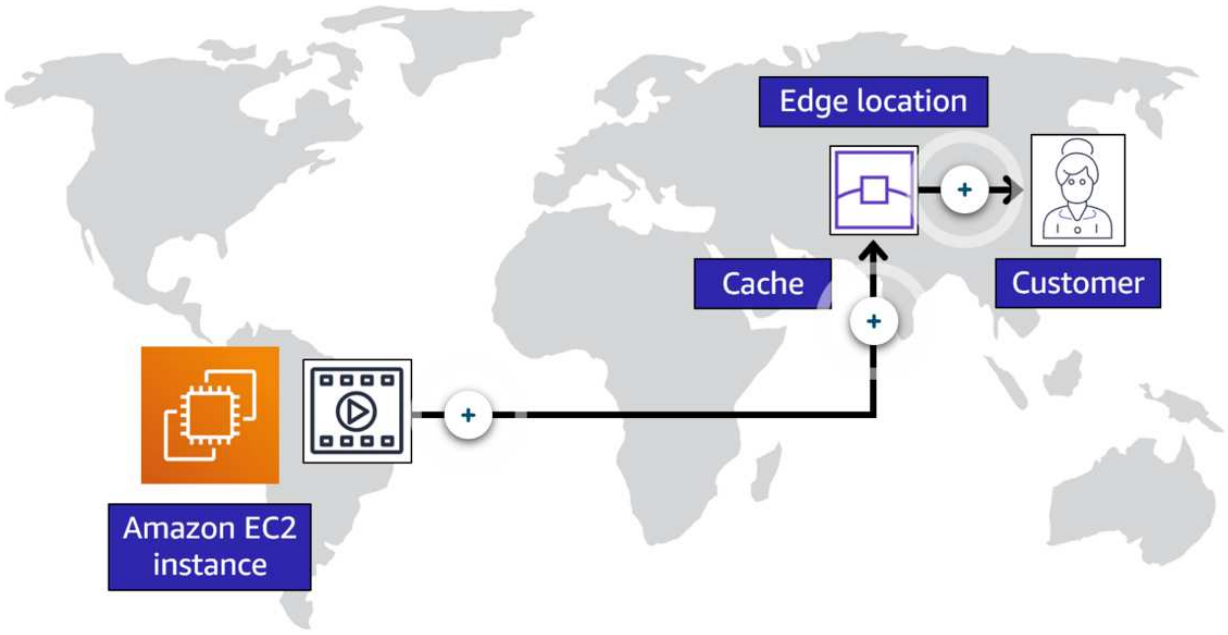

Figure 20 AWS Edge location used for Caching [lxxiii]

3.5 WHY IS AWS COMPUTE BENEFICIAL FOR THE CUSTOMER

The computer capacity provided by AWS is called Elastic Compute Cloud. It is commonly referred to as AWS EC2. It is the server service offered by AWS. The secure, resizable, and completely flexible computing capacity is available for customers. After an in-depth analysis of the EC2 service capabilities, it can be concluded it is beneficial for the customers, and it is different from the on-premises servers, as summarized and captured below -

In a traditional on-prem environment, customers-

- Pay money upfront to purchase the hardware for the servers. This cost is generally very high as customers pay upfront, irrespective of their current or future needs. This is considered to be one major expense for the customer. Customers pay large chunks of money to purchase hardware not close to their current needs. Most of the time, customers end up over or under-purchasing the hardware infrastructure, leading to significant losses and wastage of investment.
- Customers then wait for the servers to be delivered to them. Generally, this process can take days to weeks for delivery. Sometimes the delivery SLA can stretch up to months, causing unnecessary delays in projects and waste of time, effort, and cost involved in deploying IT resources on site for the deployment.
- Once delivered, these servers are installed in the customer's physical data center. The installation time can also stretch from days to months. A heavy investment is involved in deploying the IT teams for implementations, training the IT for correct deployments, and arranging all necessary software and tools for proper installations.
- The customer himself performs all necessary configurations and management on the servers. All day 0, day 1, and day n operations are performed by the customer himself. Troubleshooting and maintenance can become a major challenge for the customer if they have a siloed approach and are missing the proper tools and skill set.

In contrast, EC2 is a LAUNCH - CONNECT - USE model.

Step 1: Launch an Instance
Instances are launched using templates from a web browser. Customers can select the template as per your server requirements and configurations. Templates hold the basic configuration of instances like the operating system, application server, storage capacity, and various other

hardware configurations. Customers only need to pick and select from the various options in the templates. Templates provide a variety of choices to the customer to select and configure an instance as per their requirement. All of this is done in real-time, and there are no wait times. Hence, unlike a traditional on-prem environment, customers do not need to pay to maintain all the software and hardware options in their environment. All the options are readily available to them to pick and choose from. Another significant advantage of this approach is that customers do not have to forecast their future needs. They can add and delete resources instantly as per their current requirements. This provides high cost, flexibility, and scalability benefits for the customers. Further, since the EC2 is based on a "pay-as-you-go" pricing model, where customers pay what they consume, it provides many cost savings.

Step 2: Connect an Instance

Once configured, customers can connect the instance with just one click. There are no challenges or requirements to send IT teams to the site to perform the implementation task. Customers can connect to an EC2 instance in multiple ways. Each EC2 instance has a public and private IP to access the instance. Customers can access the instance through CLI, APIs, or a web console.

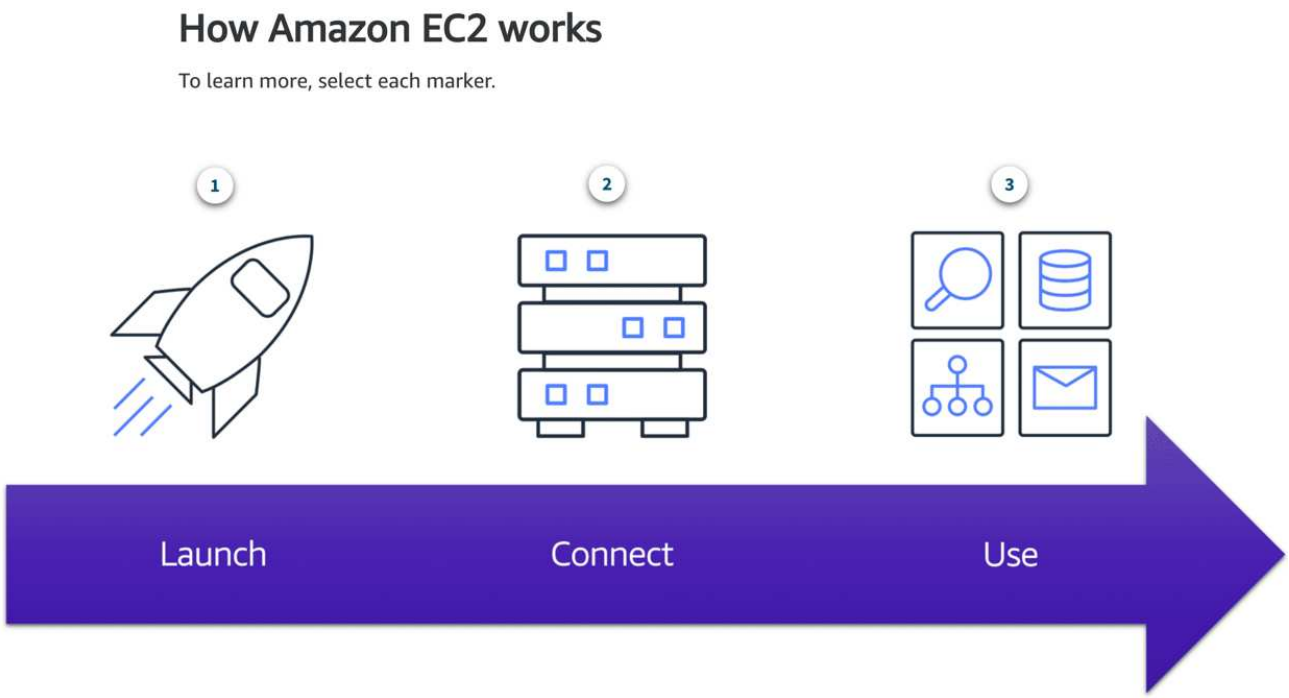

Figure 21 AWS EC2 Launch-Connect-Use model [lxxiv]

Step 3: Use an Instance

Once connected to the instance, customers can run multiple operations on the EC2 instance just like any other normal server. You can install software, add files, add storage, copy and organize data, and many more functions or perform any function without waiting hours or days for the server to be delivered, then configured and made operational and available.

Another great use case of AWS EC2 observed and analyzed was the different types of EC2 instances provided by Amazon. AWS EC2 instance types are optimized to perform many different types of tasks. When choosing an instance type, a customer must carefully think and decide about the particular requirements of their workloads and applications. This could involve demands for computing, memory, or storage resources.

Depending on these needs, EC2 provides 5 major instance types as below – [lxxv]

- General purpose instances are a balance of computing, memory, and networking resources required.
  Usage - They can be used for general-purpose application servers, small databases etc.

- Compute-optimized instances are ideal for applications that are compute-bound and require high-performance processors.
  Usage - They are generally used for compute-intensive applications like gaming servers, complex web applications.

- Memory-optimized instances are designed to provide quick performance for jobs that analyze massive datasets in memory.
  Usage -They are generally used for high-performance databases or a workload that entails handling a significant volume of unstructured data in real-time.

- Storage-optimized instances are designed for workloads that need rapid, sequential read and write access to huge datasets on local storage.
  Usage - They are generally used for distributed file systems, data warehousing applications, and high-frequency online transaction processing (OLTP) systems.

- Accelerated computing instances are generally needed for those applications which require much bigger and stronger computing powers. They use hardware accelerators or co-processors. Usage - They are generally used for graphics applications.

Amazon EC2 has an extremely flexible pricing model, where you pay only for the amount of computing you use. Studies have shown this is the biggest attraction and reason for customers to choose EC2 over their on-prem servers. There are no upfront payments or charges involved. Customer is charged what they consume, which is generally referred to as "pay-as-you-go" pricing. This flexible pricing model saves customers a lot of money and makes their networks scalable, as they do not need to plan the future server requirement. Depending on their future need, they can either increase or decrease the number of EC2 instances and pay only for the instances they use.

There are 5 most common and flexible ways to pay for EC2 instances – [lxxvi]

- On-Demand Instance
  In this model, customers pay only for the computing time they use. These instances run continuously till the time they are stopped manually. They are the most common and popular choice for small customers and induvial users. They are generally used in instances that cannot tolerate interruptions and are short-term. There are no upfront costs or minimum costs involved here.

- Savings plan
  Customers commit to a consistent compute consumption for 1 or 3 years in this model. As shared by AWS, this model can offer customers a discount of up to 72%. This model is most relevant for customers who need a consistent computing capacity for a specific period.

- Reserved Instances
  In this model, customers can reserve an on-demand instance for 1 or 3 years and get discounts. At the end of the term, customers have to pay the regular price. This model is most commonly used by customers who wish to reserve compute available for a specific period.

- Spot Instances

In this model, customers can get unused instances available at spot pricing. These prices are much lower than regular on-demand pricing. If the cost of the instance is less than the customer's bided price, the customer is assigned that instance. However, as soon as the price of the instance goes up, the customer loses that instance. They are more generally used for short-lived workloads and can handle interruptions.

- Dedicated hosts

  They are the most expensive pricing model. This is typically like the customer purchasing the entire server for their use. They can be used to maintain enterprise licenses. Enterprise organizations generally use them to hold their critical and private data in the cloud, and for security and compliance, they cannot share their computing capacity with other customers. So, they end up renting the entire server itself.

3.6 HOW CUSTOMERS GET THE MOST EFFICIENT STORAGE WITH AWS

After understanding the AWS architecture and services, it can be concluded that customers get a very efficient storage capacity with AWS because AWS provides highly scalable, reliable, extremely fast, flexible, and secure cloud storage to store and retrieve any data in the cloud. These capabilities are possible because AWS offers multiple types of storage depending on the type, usage, and capacity to meet the customer's requirements. Customers do not have to stick to one particular type of storage to enjoy these storage capacities and capabilities.

Additionally, the storage provided by AWS has many other functionalities and features which are attractive choices for customers, like encryption capabilities, auto-backup, high availability, and load balancing. All these features customers can get with a pay-as-you-go pricing model and customize the storage sizes as per their requirements.

To meet the customer requirements in terms of usage, capacity, flexibility, and pricing, the storage facilities of AWS understood and summarized below -

- Amazon Instance stores
  For an Amazon EC2 instance, an instance store provides momentary block-level storage. This type of storage is temporary and is lost once the instance is terminated. Hence, once the instance is stopped, you lose all the data stored in the instance with the instance store. This storage comes with every EC2 instance and is most commonly used for storing temporary data or testing data that is not needed in the future.

- Amazon Elastic Block Volume
  Commonly known as EBS Volume, this type of storage provides block storage, and it remains persistent even if the EC2 instance is stopped. Hence, data stored in EBS volume is not lost, even when the instance is terminated. Since EBS volume data is not lost even after the instance is terminated, it is important to back up the data in EBS. The data in EBS is backed up as EBS snapshots. EBS snapshots are not complete data backups but are incremental backups, only backing up the changes or the modifications. The incremental backup happens below

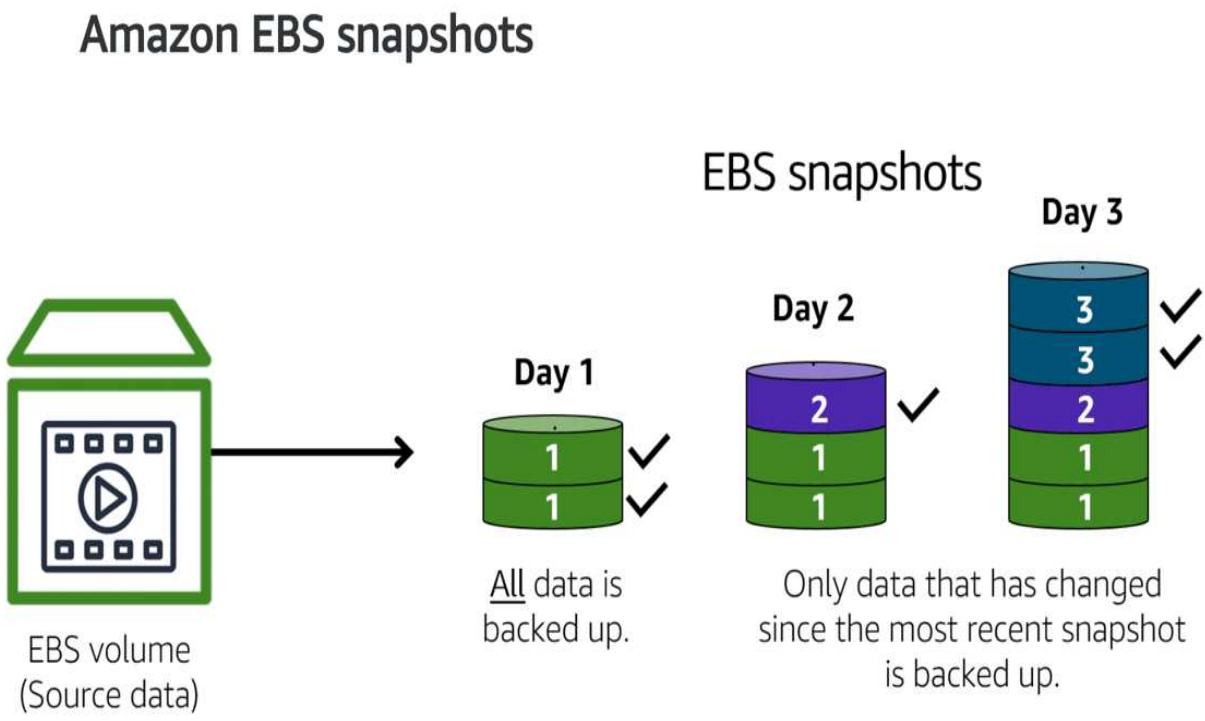

Figure 22 AWS EBS Snapshot [lxxvii]

- Amazon Simple Storage Service (Amazon S3)

Commonly known as Amazon S3, this storage class provides object-level storage. Unlike the other types of storage, this type stores data as objects, not blocks. S3 is the most common type of storage facility and is most widely used by customers. This storage is available separately, unlike EBS or instance store which is attached to EC2. Data stored in S3 is called object, and objects are stored in S3 buckets.

S3 is used to upload files like images, text, pictures, videos etc. Each object stored in the S3 bucket has three essential parts –

Object data - Object data is the data uploaded to the S3 bucket.

Object Metadata - Data about the data is called Object Metadata. It has details about the data, like the data type, size, etc.

Object Key - The object key is the unique identifier of the data.

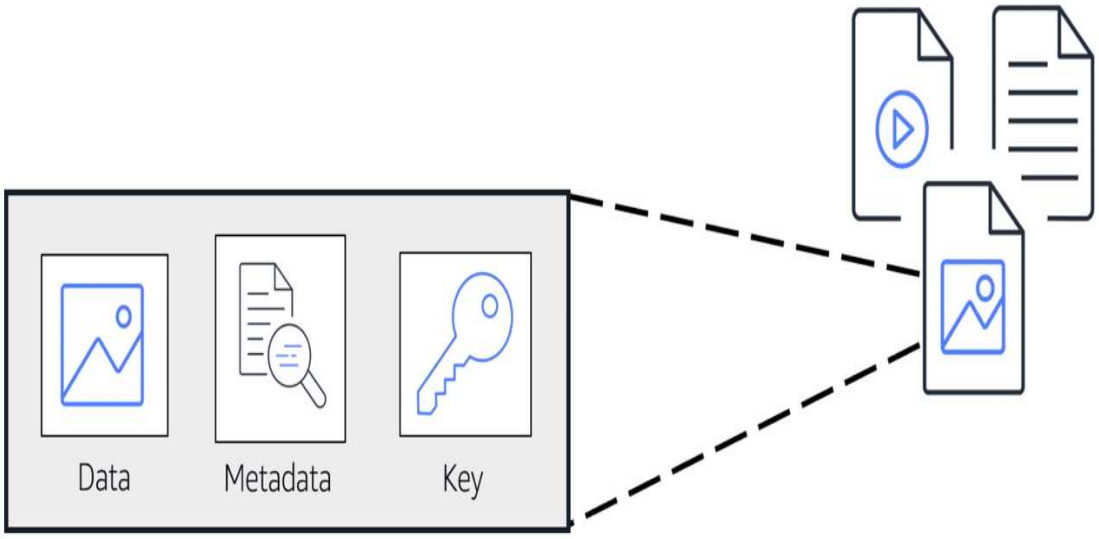

Figure 23 S3 Object [lxxviii]

Storage space in S3 is unlimited. However, the maximum size of each object size is 5TB.

Customers can control access to S3 object files by setting the permissions and enabling versioning to check the changes made to the object. Both these features have a great use case for the security and privacy of data.

Amazon S3 storage class also has “pay-as-you-go” pricing. Where you pay only for the storage consumed by you, your use. Amazon offers different types of S3 storage classes to meet different storage needs and the cost of the customer. Two important factors to consider while selecting a storage class –

- How frequently are you going to access your data?
- How accessible does your data need to be?

In order to meet different customer needs, AWS offers different types of S3 storage classes providing great flexibility in terms of storage periods, retrieval rates, and availability of the data-

Amazon S3 standard - Created for data that is frequently accessed. Data is three AZs by default.

Amazon S3 Standard-Infrequently Access- Created for data that is less frequently accessed. Helps to save costs.

Amazon S3 One Zone-Infrequently Access - Created for less frequently accessed data that does not require backup or high availability. Hence stored in just one AZ.

Amazon S3 Intelligent-Tiering - Moved data between different classes to save cost depending on how often the data is accessed.

Amazon S3 Glacier Instant Retrieval - Created for data meant for archival storage but must be accessible within a few milliseconds.

Amazon S3 Glacier Flexible Retrieval - Created for data meant for archival storage but must be accessible within flexible retrieval timings.

Amazon S3 Deep Archive - Created for data meant for archival storage but must be accessible within 12-48 hours.

3.7 WHY IT IS IMPORTANT TO USE AWS NETWORKING

A network is important to allow communication between any two types of devices. For any customer to use cloud services provided by AWS, it is very important and crucial to understand and use the networking infrastructure provided by AWS. AWS networking allows customers to design and build a fast, reliable, and secure network for their businesses and allow communication between various services rented from AWS like compute, server, storage etc. [lxxix]

Customers can divide their cloud infrastructure using AWS networking services, scale up workload requests, and connect the physical network to their private virtual networks. [lxxx]

AWS networking involves two important concepts –

- Amazon Virtual Private Cloud (VPC) – With AWS VPC, customers can create virtual boundaries between their AWS resources. For example, customers can put their critical data in one EC2 instance in one VPC and their public data in another EC2 instance in a different VPC. Amazon VPC is like a logical separation. This is important for customers to maintain the correct security and segmentation. Traffic in VPC is controlled through Network Access List. [lxxxi]

Figure 24 AWS VPC [lxxxii]

- Subnets - A subnet is a region of a VPC where resources can be grouped according to operational or security requirements. Subnets can be either private or public. A public subnet contains resources accessible by the public, whereas a private subnet can contain private and internal resources [lxxxiii] Traffic in a subnet is controlled through a security group.

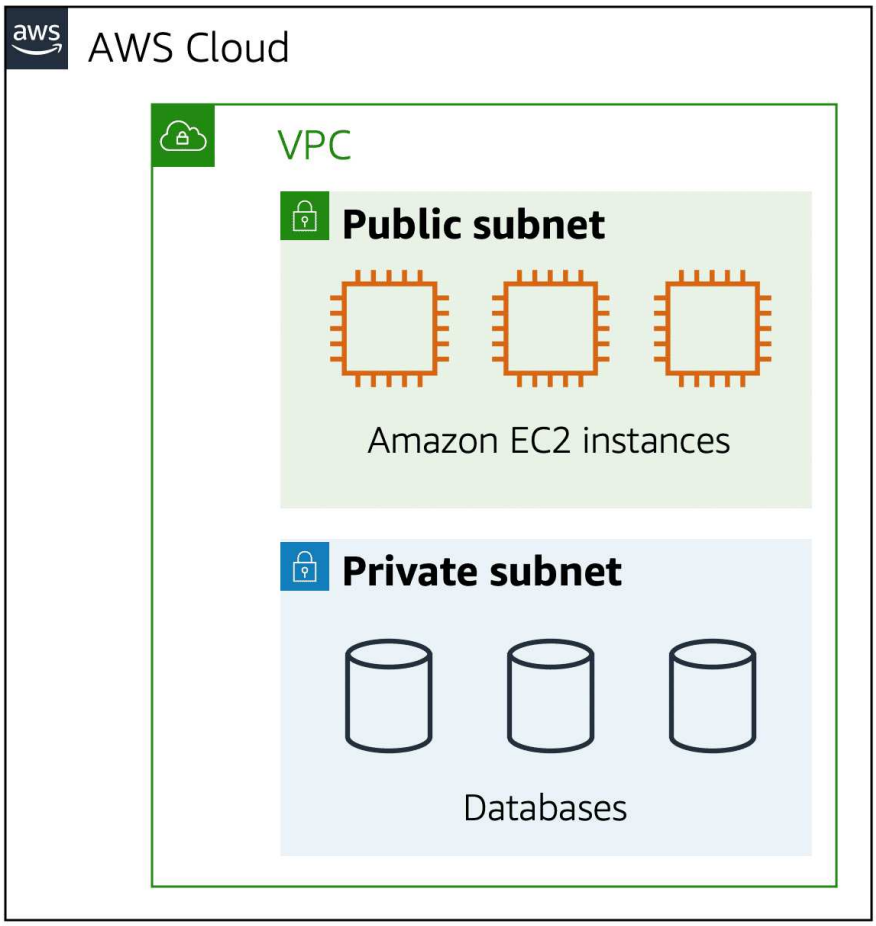

Figure 25 AWS Subnet [lxxxiv]

After studying and analyzing the networking structure provided by AWS, its importance for customers is summarized below –

- The high degree of security – AWS networking structure involving the VPC and subnets provides the necessary segmentation customers need to segment their private and public data and ensure network security.

- High network availability – With AWS networking, high-speed connectivity can be obtained between various AWS resources, even in different availability zones and regions, making the network highly available and fault tolerant.

- Consistent High Performance – AWS networking ensures minimum network downtime and fast connectivity. Because of a fully connected network, content is delivered much faster and efficiently to the end customer.

- Global Scale – AWS networking allows customers to expand their logical boundaries/segmentations across different availability zones and regions, allowing them to utilize teams and IT resources globally.

- Management & Monitoring – AWS networking can be managed, configured, and monitored through the AWS management console. The entire infrastructure architecture can be designed and configured through a single dashboard, providing ease and simplicity to the customers.

- Loading Balancing and Auto Scaling – AWS networking also provides load balancing capabilities in their networking services to manage heavy traffic to the customer network. It also has auto-scaling capabilities which can add and remove capacity according to the incoming traffic to the network.

3.8 IS AWS MONITORING EFFICIENT

After understanding the challenges of the traditional on-premise data center, it was figured out that one of the major challenges faced by customers in their traditional data centers was the lack of full visibility. Because visibility was a major issue, troubleshooting and monitoring the data was challenging. Customers spend a lot of time and effort on their IT labor and expenditure to troubleshoot in case of any error. Also, because customers lacked full visibility and monitoring capabilities, the data centers were vulnerable to security breaches and major attacks on the internet. Often, security breach incidents would happen because of the internal employees themselves because there were no strict control and observability measures in place. Another challenge customers faced by customers due to lack of visibility was their inability to innovate. Because customers had limited understanding and observability of their environment, they could not innovate efficiently. They were also incapable of deploying new methods and tools to drive growth and innovation in their business.

In contrast, the AWS cloud services model is helping customers by providing full-stack observability. Customers can enable, provision, and manage their environment with AWS for governance, control, and business agility. [lxxxv]

AWS provides different tools and methods for full stack observability through a single dashboard, i.e., the AWS management console. After understanding customer adoption, some of the most valuable and highly rated management services provided by AWS are summarized below –

- Amazon Cloudwatch - To make managing your infrastructure and applications more efficient, Amazon CloudWatch gathers and displays real-time logs, metrics, and event data in automated dashboards. [lxxxvi] AWS cloud watch is used by customers for monitoring application performance, performing root cause analysis, optimizing resources proactively, and testing website impacts. [lxxxvii] Amazon cloud watch can also be used to set alarms and alerts in case of any consumption within the customer's network goes up or below the set threshold.

Figure 26 AWS Cloudwatch [lxxxviii]

- Amazon CloudTrial - To give you control over storage, analysis, and corrective measures, Amazon CloudTrail keeps track of and logs account activity across your AWS infrastructure. AWS Cloudtrail gives customers answers for the who, what, when, and where of any action in AWS cloud. [lxxxix]

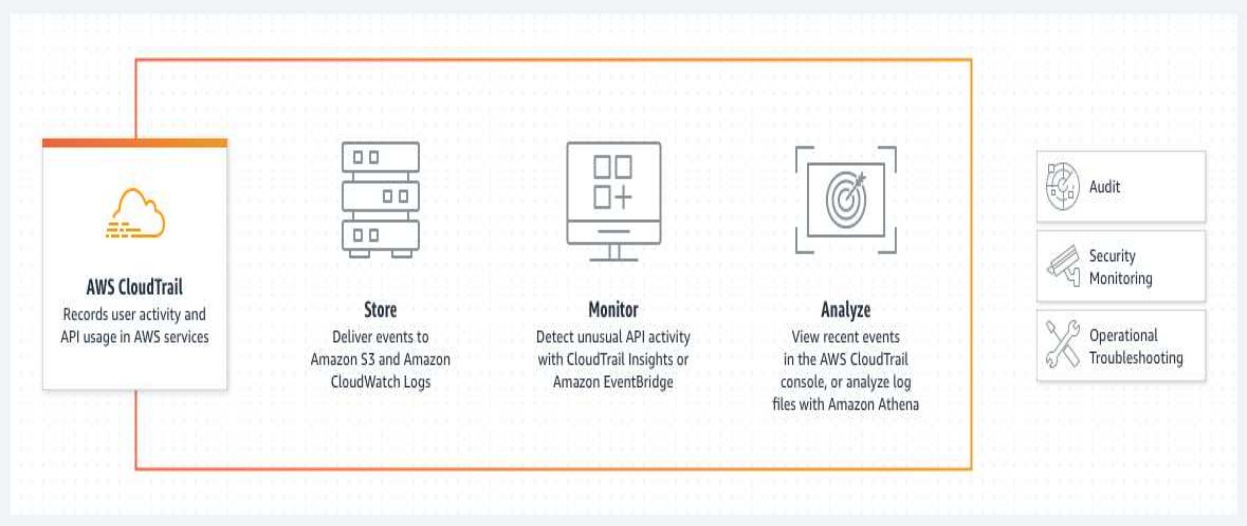

Figure 27 AWS CloudTrail [xc]

- AWS Trusted Advisor – AWS trusted advisor provides real-time recommendations and suggestions to customers to change and modify their settings based on the 5 pillars of AWS Best Practices – Cost optimization, Performance, Security, Fault Tolerance, and Service Limits. [xci] The trusted advisor helps customers to optimize their infrastructure, improve security, compliance, and performance, reduce costs and expenditure, and continuously monitor service quotas.

Figure 28 AWS TurstedAdvisor Dashboard [xcii]

3.9 HOW TO MIGRATE ON-PREMISES APPLICATIONS TO AWS CLOUD

Because AWS cloud services offer many features and capabilities to customers, more and more organizations are moving towards adopting the cloud and moving their workloads to the cloud. Moving their applications to the cloud provides customers more flexibility, scalability, reduction in cost benefits, agility, and time to innovate faster.

However, cloud migration is a complex and detailed process. It has been observed that most customers fail to enjoy the full benefits of cloud services because of their poor migration strategies. [xciii] Before migrating to the cloud, customers need to design a comprehensive strategy considering several possible failure issues like security gaps, interoperability challenges, lack of planning, unplanned expenditures, lack of understanding of the cloud infrastructure, and poorly designed architectures. [xciv]

After studying AWS's challenges and infrastructure capabilities, the below 6 strategies are considered vital for any cloud migration. Every customer looking to adopt the cloud must adopt one or more of the below strategies to ensure a successful migration. These 6 strategies and their customer relevance have been summarized below-

- Rehosting

  Rehosting is also called a 'lift-and-shift' mechanism. [xcv] This is a legacy method of migration, which involves migrating a workload to the AWS cloud without any changes or optimization. This is particularly useful for customers who quickly want to move to the cloud and start enjoying the benefits. Their workloads are not so complex, and they can quickly adopt the cloud.

- Replatforming

  Replatforming is also called the "lift, tinker, and shift" mechanism. [xcvi] This migration method involves optimizations and changes in the workload before moving it to the cloud. These small changes allow customers to enjoy the full benefits of cloud migration. However, the core of the architecture is not changed or altered in the process. A prevalent example of this strategy is adding a few licenses to customer applications to make them operational on the cloud, like authorization licenses. And then migrating the applications to the cloud without changing the core architecture of the application.

- Refactoring (also known as re-architecting)

  In the refactoring strategy, the application's architecture needs to be changed to make it optimized and interoperable on the cloud and to use cloud-native features. It is suitable for those customer applications whose architecture cannot be deployed in the cloud and needs rearchitecting to match the features in the cloud.

- Repurchasing

  The repurchasing strategy basically requires customers to purchase something new. In other words, move to a new or a different product to use cloud-native features. A common example of this strategy is shifting a CRM to Salesforce.com, an HR system to Workday etc. [xcvii]

- Retaining

  In retaining customers, they are advised not to migrate their workloads to the cloud and keep them in their original environment. Because these applications are critical to the customer, migrating them to the cloud needs a full re-vamp and re-architecture. Sometimes re-vamping is also impossible due to other reasons like expenditure, labor, and risk factors.

- Retiring

  Under retiring, customers are advised to remove the applications as they are not needed anymore and not to move them to the cloud.

3.10 ANALYSIS OF AWS CLOUD AGAINST OTHER CLOUD VENDORS

AWS is the biggest cloud vendor in the market. While it is ahead of other small cloud vendors, its two main competitors are Google Cloud Platform and Microsoft Azure. AWS cloud is the market leader in cloud services, followed by Microsoft Azure and the Google Cloud Platform.

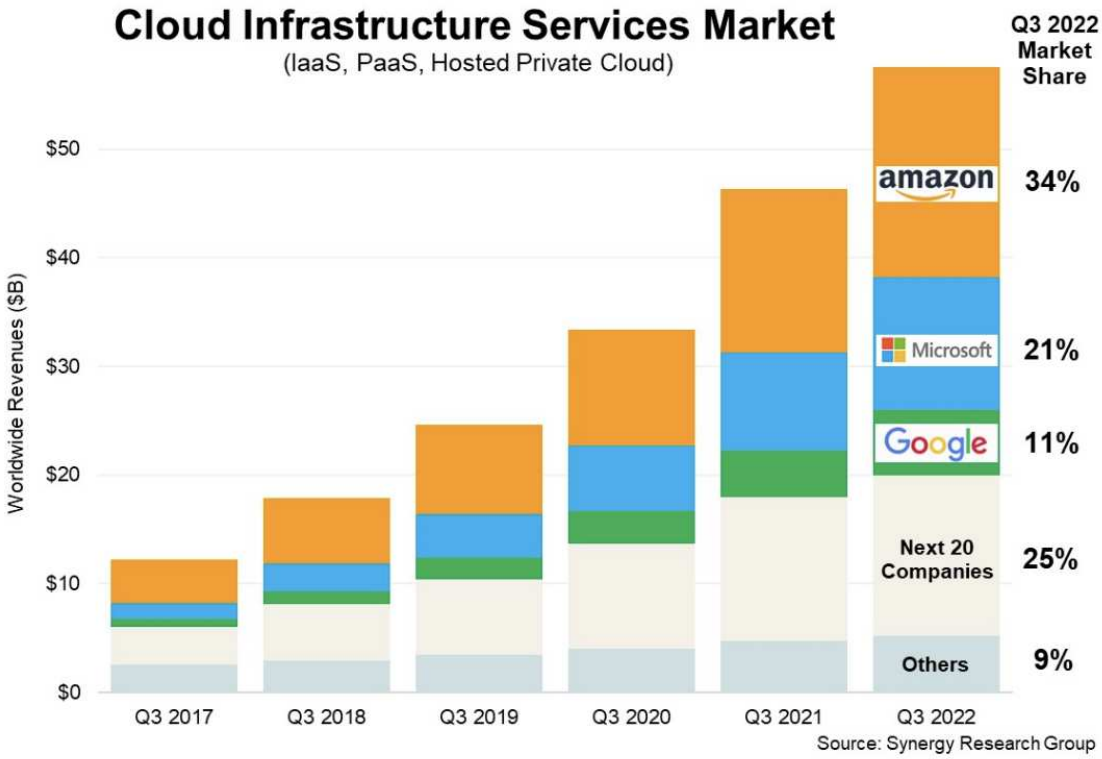

Figure 29 Market Share of AWS, Azure, and GCP [xcviii]

Even though AWS is the market leader in cloud services, Azure is rapidly becoming famous with enterprise customers, especially with its Microsoft teams solution, and Google Cloud is also continuing to evolve rapidly. After a thorough analysis, the key differences between the three big cloud vendors are summarized below –

- Region and Availability

Currently, AWS has 31 Regions and 99 Availability zones. Azure has 60+ announced regions, and GCP runs in 27 cloud regions. Azure has the broadest reach and coverage in terms of the region compared to the other two vendors. [xcix]

- Common Services – Both AWS and Azure provide 200+ cloud services. GCP, on the other, provides nearly 100 + cloud services. The broadest range of services is offered by AWS. With an excellent array of analytics, AI, and machine learning offerings, Azure is a close second. Due to the limited number of services provided, Google Cloud Platform sits in third place. [c] All three vendors provide comparable computing, storage, and networking services, as shown below –

Compute Services

SERVICE	AWS	AZURE	GCP
VM (Compute Instance)	EC2 (Elastic Compute)	Azure Virtual Machine	Google Compute Engine
PaaS	AWS Elastic Beanstalk	App Service	Google App Engine
Container	AWS Elastic Container/Kubernetes Service	Azure Kubernetes Service (AKS)	Google Kubernetes Engine
Serverless Functions	AWS Lambda	Azure Function	Google Cloud Functions

Figure 30 Comparison of Compute Services by Aws, Azure and CGP [ci]

Database & Storage Services

SERVICE	AWS	AZURE	GCP
RDBMS (Multiple Database Types – SQL, MySQL, etc..)	AWS RDS	Azure SQL/ Database for MySQL/PostgreSQL	Cloud SQL
NoSQL	DynamoDB, Simple DB	Azure Cosmos DB, Table Storage	BigTable, Cloud Datastore
Object Storage	S3 (Simple Storage Service)	Blob Storage	Google Cloud Storage
File Storage	Elastic File System	Azure File Storage	Google Filestore
Archive Storage	Amazon Glacier	Azure Archive Storage	Google Storage (Archive Storage)
Data Warehouse/Data Lake	Amazon Redshift	Azure Synapse Analytics	Google BigQuery

Figure 31 Comparison of Database services by AWS, Azure, and GCP [cii]

All three vendors have a strong portfolio of computing, Networking, and Storage services. The major difference which makes a distinction between these services is the architecture of these services in the cloud and how easily the induvial features of each service are available to the customers for consumption. As per the market analysis, AWS features to provide maximum features, capabilities, and flexibility. Hence, it is the most widely used.

Networking

SERVICE	AWS	AZURE	GCP
Virtual Network	Virtual Private Cloud (VPC)	Virtual Network (Vnet)	Virtual Private Cloud (VPC)
Load Balancing	Elastic Load Balancer	Azure Load Balancer	Google Cloud Load Balancing
Firewall	AWS Firewall / Web Application Firewall	Azure Firewall	Google Cloud firewalls
DNS	Route 53	Azure DNS	Google Cloud DNS
CDN	Amazon CloudFront	Azure Content Delivery Network (CDN)	Cloud CDN

Figure 32 Comparison of Networking services by AWS, Azure, and GCP [ciii]

- Pricing – It is difficult to comment on the pricing of the services provided by three cloud providers. All of them provide competitive pricing for all their services. However, the underlying pricing model is the same for all the vendors: subscription-based and pay-as-you-go. The price of service any vendor provides depends largely on customer-specific requirements, amount of usage, and business needs. AWS provides a pricing calculator which helps customers estimate their usage, as their pricing can get a little complex. Azure pricing is considered much more straightforward, with no hidden costs. GCP is a leader in this domain, offering the most straightforward and customer-friendly pricing.

CHAPTER 4

PROJECT IMPLEMENTATION

4.1 AGENDA

The agenda of the implementation is to migrate an on-prem application to the AWS cloud while understanding the necessary components and architecture of the AWS cloud. This implementation will help understand how cloud computing can overcome the challenges of on-premises systems. The implementation process will provide an opportunity to learn the principles of Cloud computing, compare and contrast AWS cloud computing with on-premise models and understand the details of various AWS architecture, components and services.

4.2 METHODOLOGY

To perform this implementation, below resources were used –

- Software Requirement –AWS account to access the AWS console and services, Oracle VM Virtual Box, LINUX, APACHE, SQL, PHP WordPress Packages,
- Hardware Requirement – Laptop to host and run the applications and a VM with a working terminal
- Key Source of Learning - AWS cloud networking learning library, AWS Skill Builder

To showcase this migration using cloud networking, a sample test application was first hosted locally on the laptop. In order to create a sample application, a Virtual Machine was deployed on the laptop with Ubuntu as the operating system; for this Oracle VM virtual box was downloaded. The sample test application hosted on this VM was a WordPress application

4.2.1 RATIONALE OF CHOOSING WORDPRESS AS THE INTENTED WEB-BASED APPLICATION

For this project, WordPress was chosen as a sample test application. This is because of it's wide popularity. It is the world's most popular content management system and more than 34% of the websites on the internet today run on wordpress. [civ] More than 500 websites are build each day using WordPress. [cv] It is such a popular choice among the customers because it is free, supports open source, easy to use, provides endless customization options and easy to manage. Since, WordPress is so widely used and it made a good choice for the sample test application. Most of the of the web applications will have a similar web application architecture and this migration strategy can be replicated for the them as well. Hence, considering the popularity and user acceptance rate of WordPress, it was chosen as a sample test application for this project.

WordPres is also based on the most traditional LAMP stack. Since, WordPress is based on LAMP stack, each of its components like web server, front-end application code, support framework like PHP, and back-end SQL data base, all can have a clear path of migration to cloud. LAMP stack is the used by most of the developers to create, host and maintain web content. Almost all websites on the internet are based on the LAMP architecture and it provides a streamlined approach for cloud migration for web applications.

4.2.2 WHAT IS A LAMP ACHITECTURE?

A LAMP stack is a set of four software packages, which are used by developers to build and host websites. These four software technologies include – LINUX, APACHE, MYSQL, and PHP.
Linux is the operating system, APACHE is the web server, MYSQL is for the database and database server and PHP is the programming language to create the website. [cvi]All of these software technologies are open-source and can be used by anyone freely. LAMP stack can be used to make both dynamic and static web content for the applications.

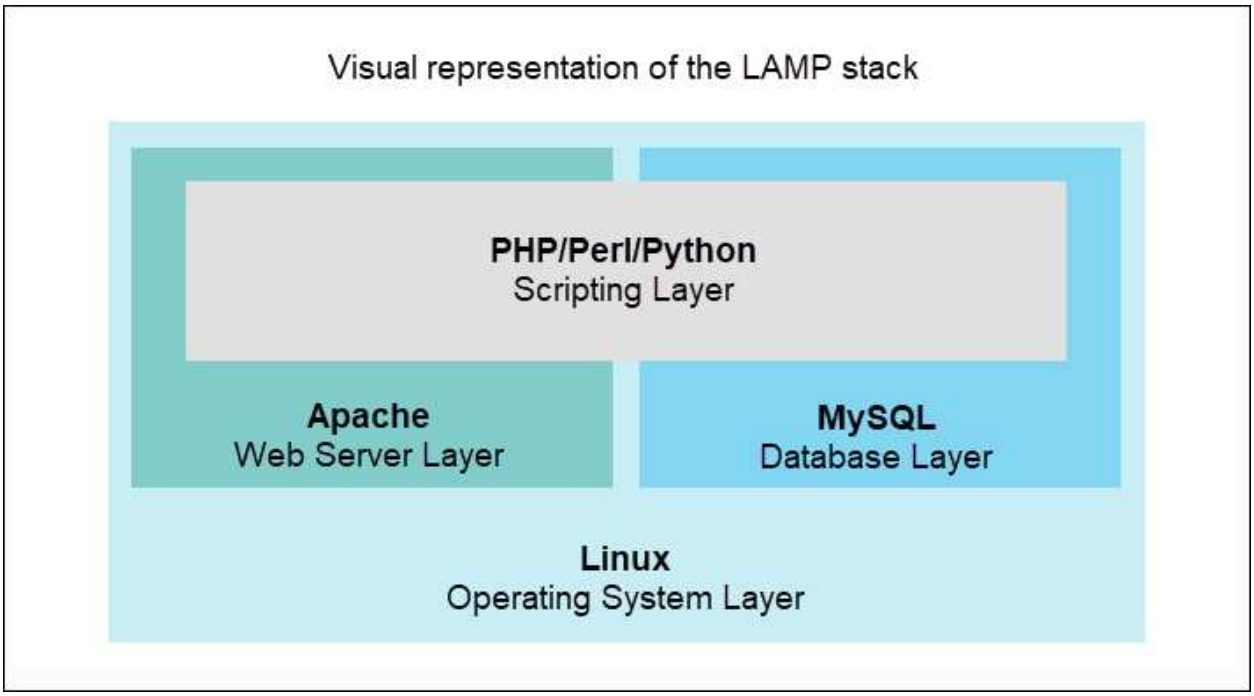

Figure 33 LAMP Stack Architecture [cvii]

These components of the LAMP architecture work together to deliver the web content of the applications as explained below – [cviii]

1. The Apache web server and MySQL database run on the Linux operating system and communicate using PHP. [cix]
2. Users navigate to a website and requesst on the web browser to load the contents of the website.
3. If the request is for static content, the APACHE webserver presents that directly.
4. If the user request is dynamic content and a PHP file type, Apache uses PHP to execute the request and load the contents.
5. However, if the request is not PHP type and is other data requests, then Apache fetches MySQL data.
6. PHP and MySQL provide requested resources to the Apache web server.
7. Apache webserver processes these resources to present them as HTML content to the end users.

4.2.3 WHY IS A LAMP STACK IMPORTANT FOR WEB APPLICATIONS ?

LAMP stack is most widely used by developers due to the following reasons –

- Cost – The LAMP stack is open source and the softwares in this stack are free for anyone to download and use. This is in contrast to paid and proprietary software which needs licenses for use. [cx]

- Efficiency – LAMP stack is a traditional and widely used method for web development and it is already tried and tested. All the software technologies in the LAMP stack work well in conjunction with each other and are compatible with each. Developers can start using it directly instead of worrying about their compatibility.

- Maintenance – Software experts across the globe work regularly to maintain and update the source code of these open-source software by reviewing, commenting and changing the necessary code portions. [cxi]

- Support - LAMP stack has a sizable, international community which provides support for websites using LAMP stack. Web developers can consult sample codes or make use of open-source plugins that have been well-vetted. [cxii]

- Flexibility – While the LAMP stack provides a set of four software technologies, it provides developers with a lot of flexibility as well. For example, developers can also select any other operating system other than Linxus if needed. [cxiii]

- Reliability – The LAMP stack provides a lot of reliability and resilience to the websites to its flexible, open source and end-to-end architecture for websites.

4.2.4 STEPS OF MIGRATION TO AWS CLOUD

Step 1: Download the Ubuntu Image. This image was downloaded using the below link - https://ubuntu.com/desktop

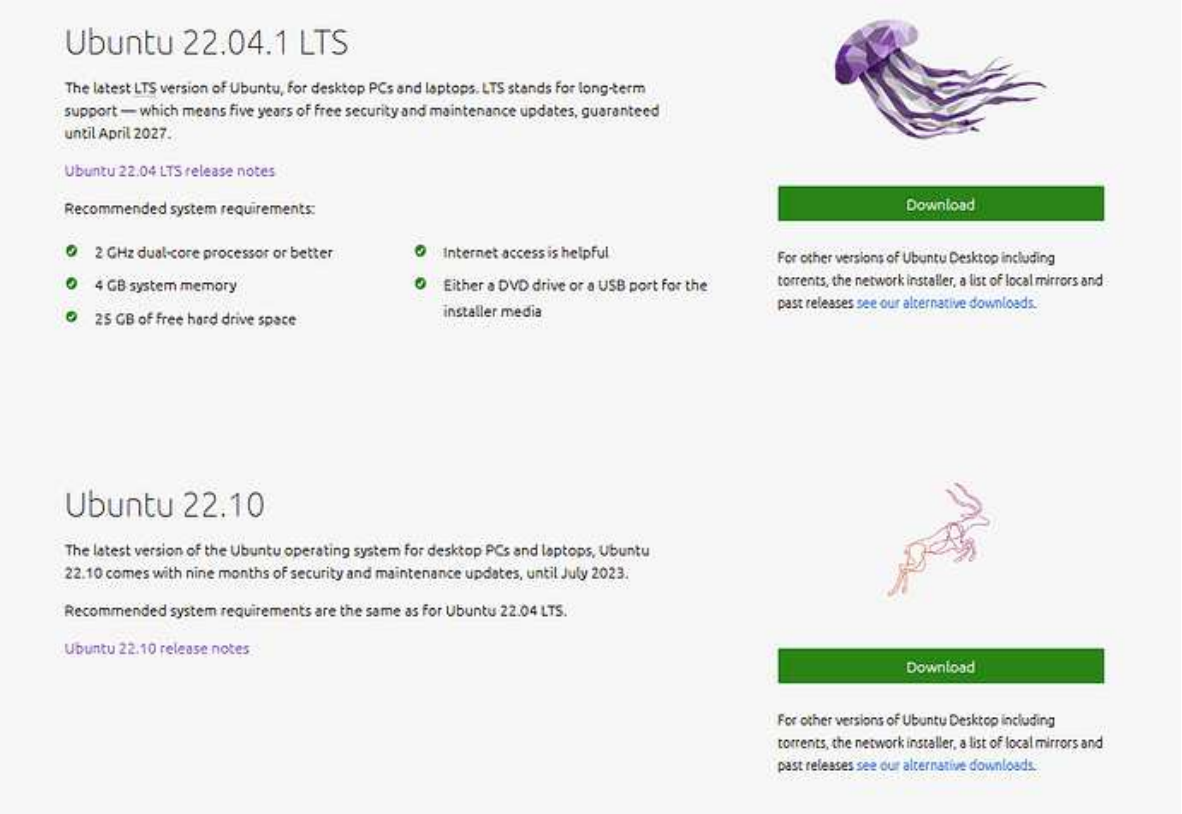

Figure 34 Ubuntu Image

Step 2: Download and install VirtualBox. This software was downloaded using the below link- https://www.virtualbox.org/wiki/Downloads

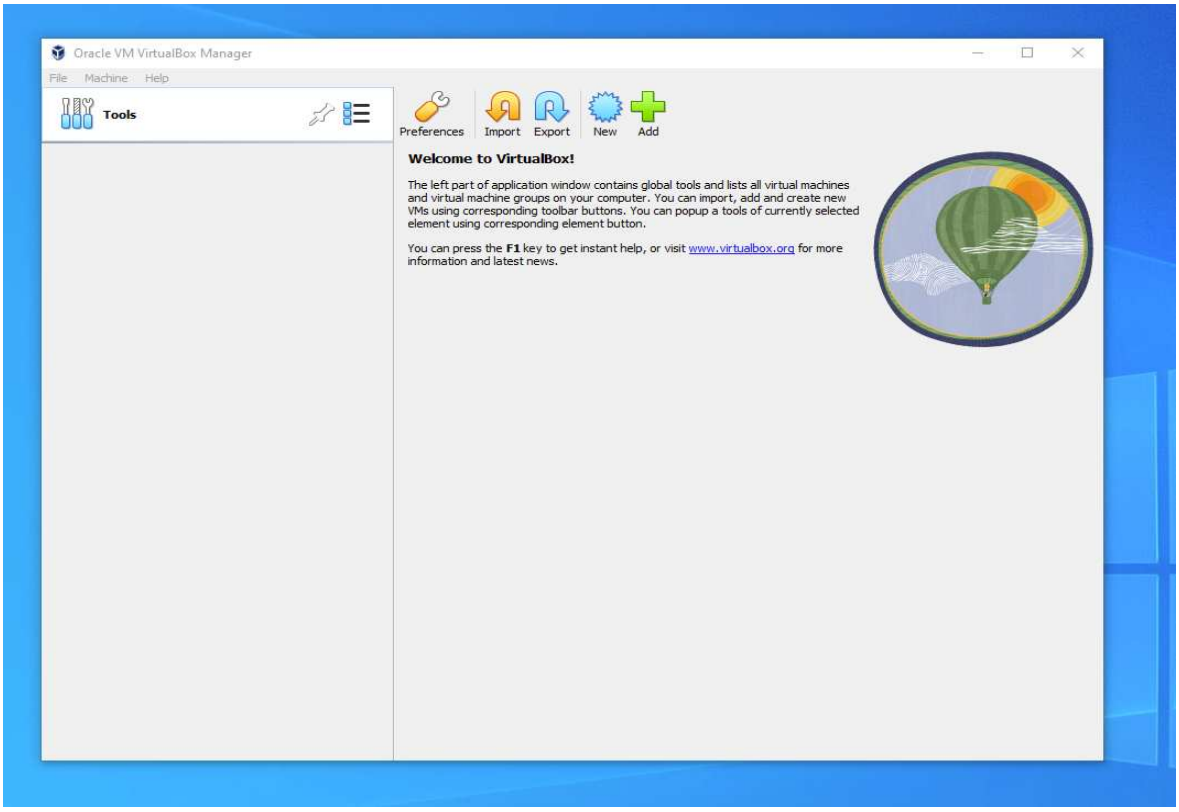

Figure 35 Oracle Virtual Box

Step 3: Ran the Virtual Box and Created a new virtual machine. Field details:

- Name: Ubuntu in your name the Type and Version will auto-update.
- Machine Folder: Location of the virtual machine
- ISO Image: Add the Ubuntu image downloaded
- User profile: Default name vboxuser, password: 'need to be assigned.'
- Define the VM resources: 8GB RAM, 4 CPUs
- Hard disk: 25 GB as a minimum

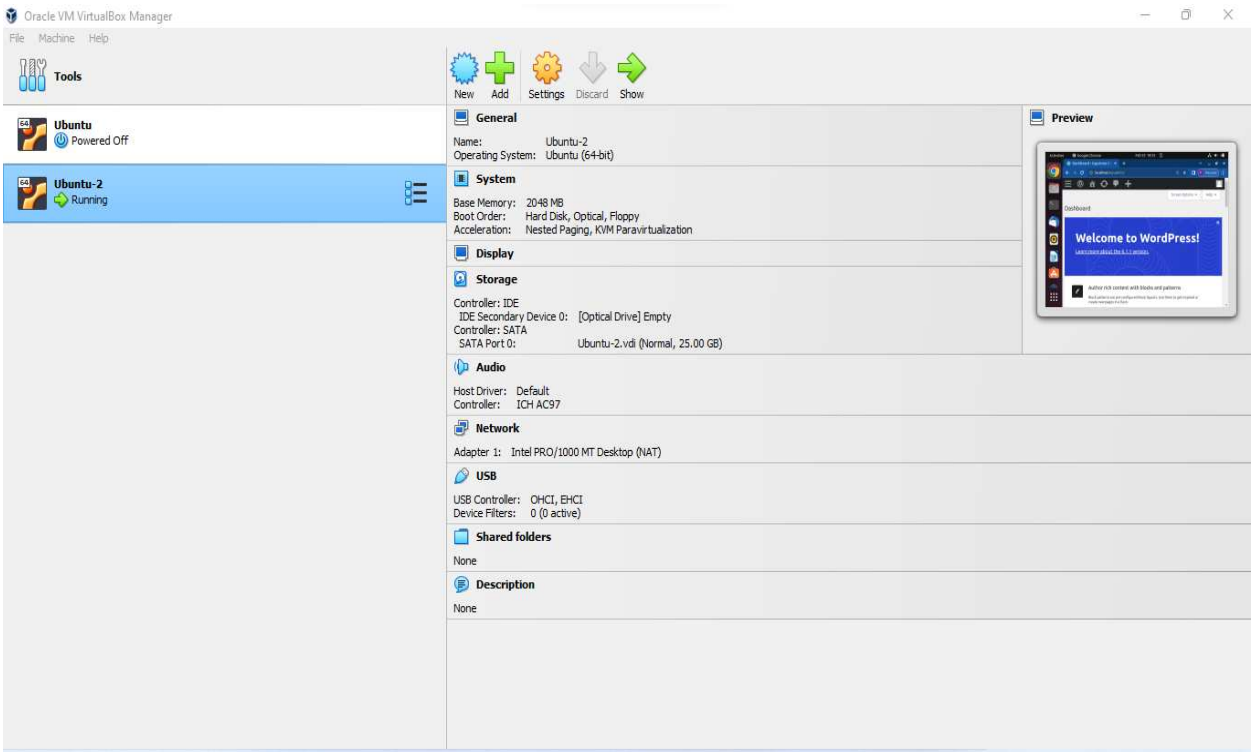

Figure 36 Virtual Machine created on VirtualBox

Step 4: Started the Virtual Machine with Ubuntu image.

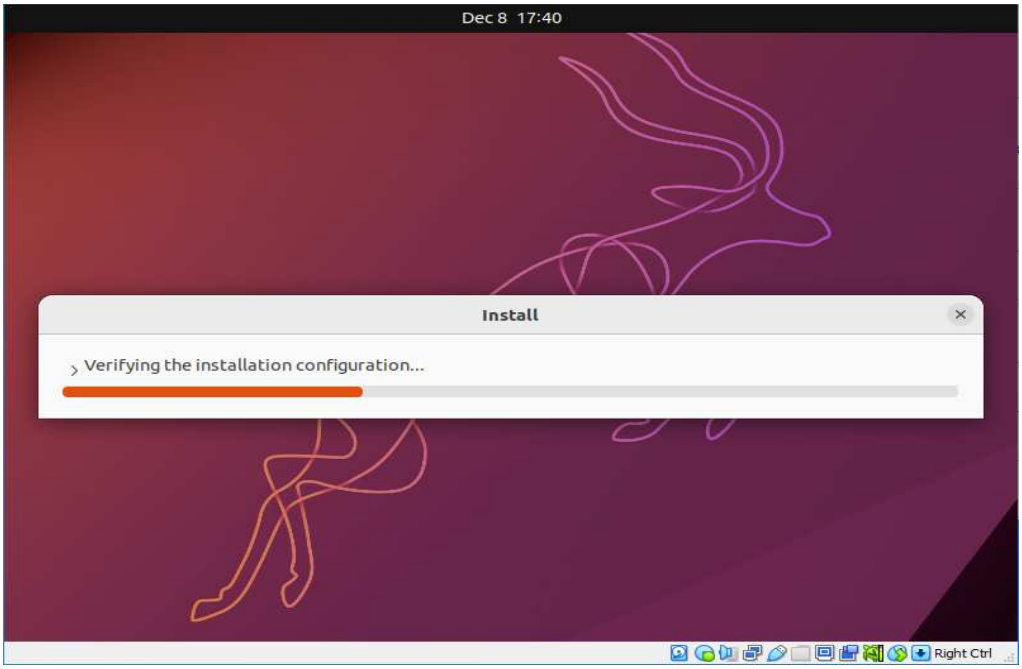

Figure 37 Ubuntu Image installation on VM

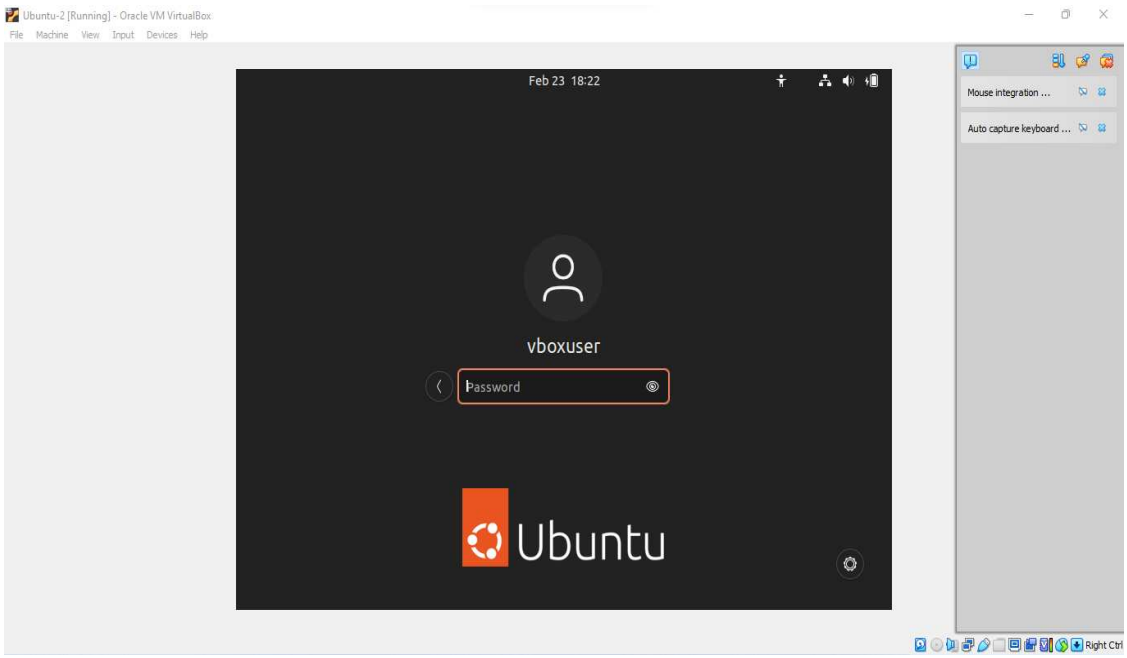

Figure 38 VM with Ubuntu OS created

Step 5: Ran WordPress locally on this VM using the WordPress package and understood the LAMP stack used.

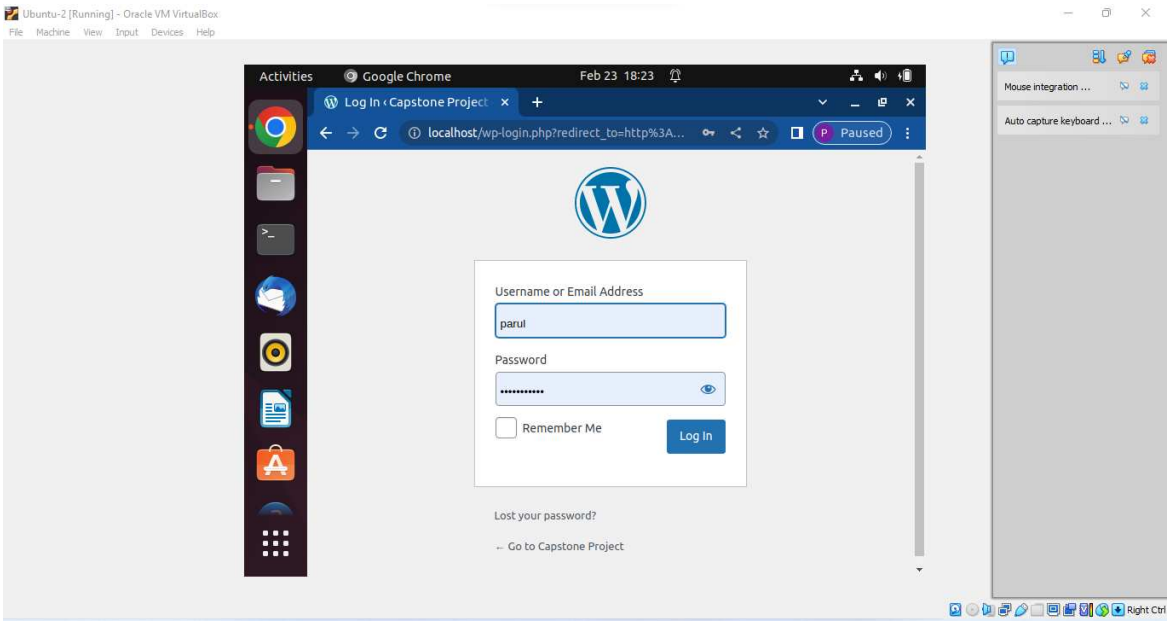

Figure 39 WordPress application created locally

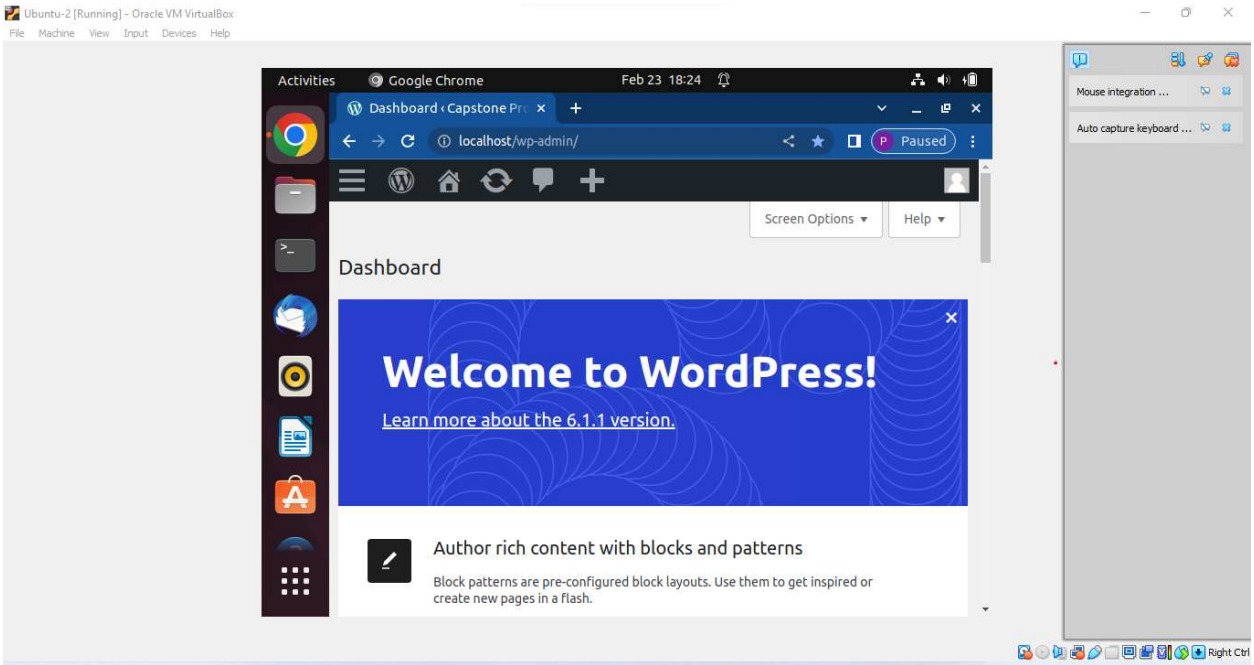

Figure 40 Wordpress Application

Step 6: Created an AWS account on the AWS management console - https://aws.amazon.com/

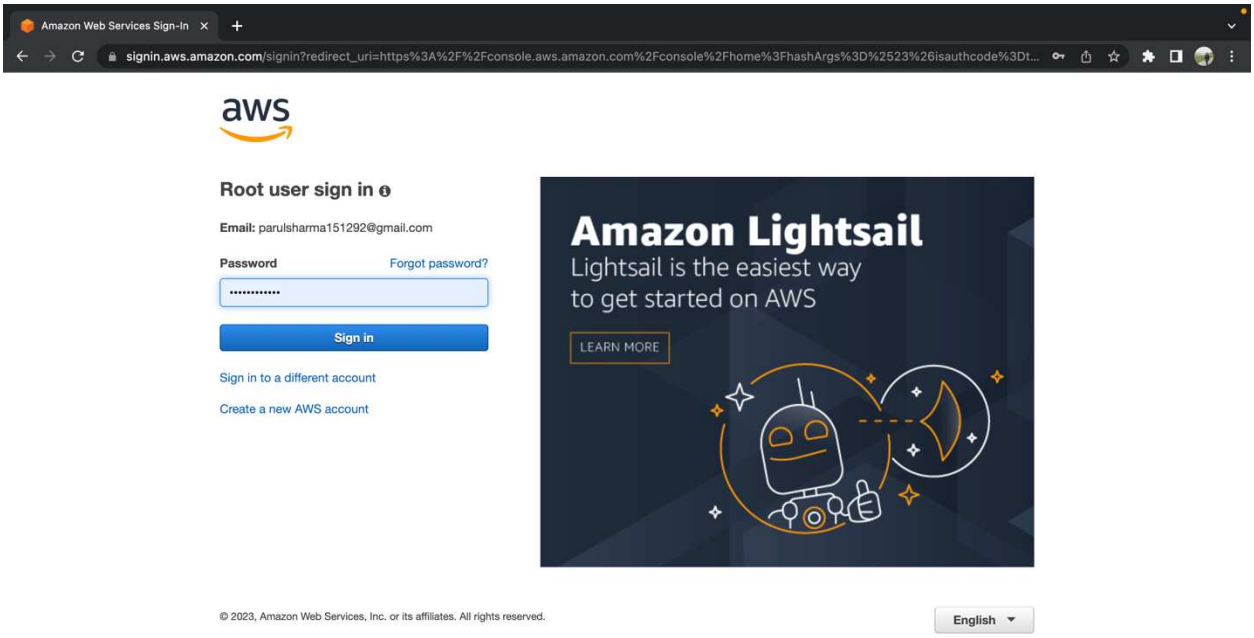

Figure 41 AWS management console to sign-in

Step 7: Design and configuration of different AWS components to migrate the web application

- Selected EC2 from the 'All services' menu and created an EC2 instance.

- Instance Name: CapstoneProject

- Server Image– Ubuntu, 22.04 LTS, amd64 jammy image buildThis Ubuntu image is available in the free tier and was selected as the original web application hosted on the local server was on Ubuntu OS.

- Instance type – T2 micro
  This instance type was selected because is available under the free tier and provides one 1vCPU, 1 GiB of memory and ESB storage. This capacity is enough to support the network performance of the WordPress application. However, if the application is complex and large bigger instance types should be selected.

- Key pair – NewKeyforCP.pem
  A new key pair was created and stored locally on the machine. This key pair is needed to perform authentication.

- VPC configuration
  The website was hosted on the us-east-1e availability zone and was provided with a default VPC - 172.31.0.0/16 and a default subnet within this VPC. The public and private IP of the application is provided from this subnet and it serves as a logical boundary for this website in the AWS cloud.

- Security Group
  The security group created for this application allowed the below connections –

Protocol	Port Range	Source
SSH	22	0.0.0.0/0
HTTP	80	0.0.0.0/0
HTTPS	443	0.0.0.0/0

- Remaining settings – Kept as default

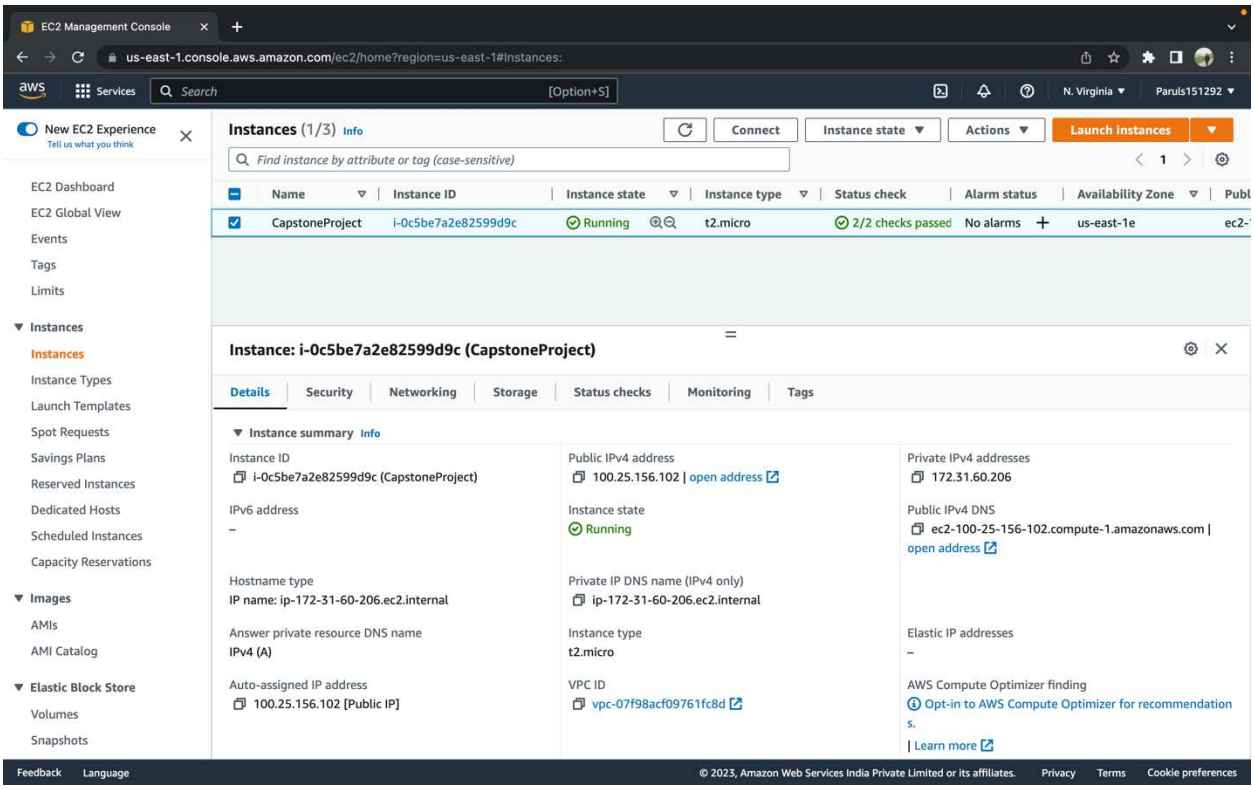

Figure 42 EC2 instance for the application

Step 8: Executed the below script to EC2 instance SSH console to migrate the application to the AWS cloud.

 # installed PHP and Apache

```
sudo apt update
sudo apt install apache2
sudo apt  ghostscript
sudo apt  libapache2-mod-php
sudo apt mysql-server
sudo apt php
sudo apt php-bcmath
sudo apt php-curl
sudo apt  php-imagick
sudo apt  php-intl
sudo apt php-json
sudo apt  php-mbstring
sudo apt  php-mysql
sudo apt  php-xml
sudo apt php-zip
```

 # created an installation directory and downloaded wordpress package

```
sudo mkdir -p /srv/www
sudo chown www-data: /srv/www
curl https://wordpress.org/latest.tar.gz | sudo -u www-data tar
zx -C /srv/www
```

 # configured Apache for wordpress

nano /etc/apache2/sites-available/wordpress.conf

```
<VirtualHost *:80>
    DocumentRoot /srv/www/wordpress
    <Directory /srv/www/wordpress>
        Options FollowSymLinks
        AllowOverride Limit Options FileInfo
        DirectoryIndex index.php
        Require all granted
```

```
    </Directory>
    <Directory /srv/www/wordpress/wp-content>
        Options FollowSymLinks
        Require all granted
    </Directory>
</VirtualHost>
```

#site enabled

```
sudo a2ensite wordpress
```

#URL rewrite enabled

```
sudo a2enmod rewrite
```

#reloaded Apache to apply the configuration changes

```
sudo service apache2 reload
```

#data base configuration

```
mysql> CREATE DATABASE wordpress;
mysql> CREATE USER wordpress@localhost IDENTIFIED BY
'<parulsh1512>';
mysql> GRANT SELECT,INSERT,UPDATE,DELETE,CREATE,DROP,ALTER
    -> ON wordpress.*
    -> TO wordpress@localhost;
mysql> FLUSH PRIVILEGES;
mysql> quit
```

#enabled MYSQL

```
sudo service mysql start
```

#connected the databased to the wordpress site

```
sudo -u www-data cp /srv/www/wordpress/wp-config-sample.php
/srv/www/wordpress/wp-config.php
sudo -u www-data sed -i 's/database_name_here/wordpress/'
/srv/www/wordpress/wp-config.php
```

```
sudo -u www-data sed -i 's/username_here/wordpress/'
/srv/www/wordpress/wp-config.php
sudo -u www-data sed -i 's/password_here/<parulsh>/'
/srv/www/wordpress/wp-config.php
```

#In a terminal session open the configuration file in nano:

```
sudo -u www-data nano /srv/www/wordpress/wp-config.php
```

Once data based configuration file is opened in nano and Find the following:

```
define( 'AUTH_KEY',         'put your unique phrase here' );
define( 'SECURE_AUTH_KEY',  'put your unique phrase here' );
define( 'LOGGED_IN_KEY',    'put your unique phrase here' );
define( 'NONCE_KEY',        'put your unique phrase here' );
define( 'AUTH_SALT',        'put your unique phrase here' );
define( 'SECURE_AUTH_SALT', 'put your unique phrase here' );
define( 'LOGGED_IN_SALT',   'put your unique phrase here' );
define( 'NONCE_SALT',       'put your unique phrase here' );
```

 # Delete these line and replace with content of https://api.wordpress.org/secret-key/1.1/salt/

 # This step is performed to ensure that site is not vulnerable to "known secrets" attacks. This just to add random values to the data base and for site protection.

4.3 OUTPUT

4.3.1 WEBSITE MIGRATED

Connected to the EC2 instance using the public IP and the WordPress application was loaded.

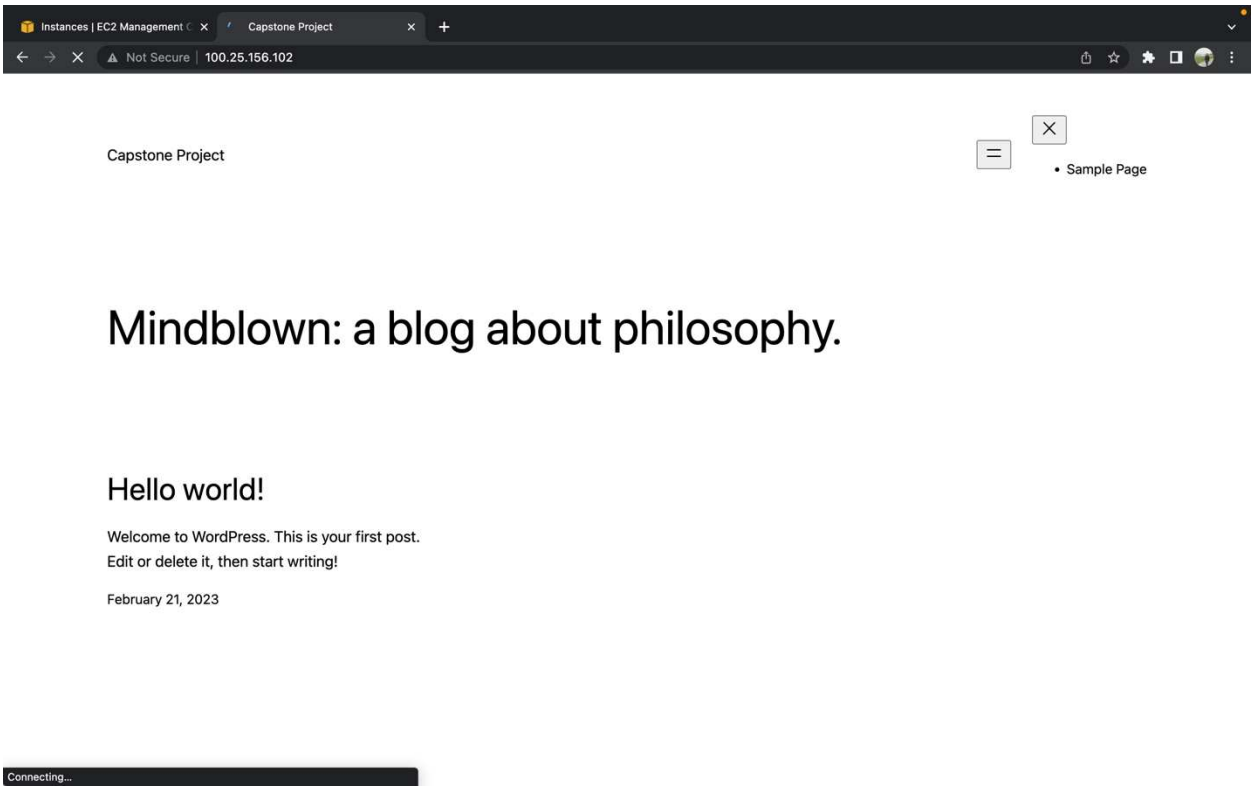

Figure 43 Wordpress on AWS cloud

4.3.2 BENEFITS OF THIS MIGRATION

Migrating this locally hosted web application to AWS cloud, provided several benefits as summarized and observed below –

- Performance – When the website content and data were hosted locally, the performance of the application was guaranteed. When the user who was sitting far away from the actual website server tried to access the website, he observed a lot of latency, as compared to the user sitting closer to the web server. In contrast to this, in AWS architecture due to a global content delivery system and caching capabilities at the edge location, the latency is significantly reduced and users' performance remains the same from any part of the world

- Reliability – The reliability and availability of the web application is automatically increased in the AWS cloud, as High availability of different components can be provided. HA can also be provided across AZs and regions to provide more resiliency. In contrast to this, in a locally hosted website, achieving such resiliency is a complex and expensive process.

- Scalable – The website hosted in the AWS cloud is scalable with default Auto-Scalaing and Elastic-Load Balancing (EBS). With Auto-scaling, AWS cloud adds more instances to the website if there are a huge number of requests coming to the website to meet the capacity. With EBS these incoming requests are spread across different EC2 instances to prevent one instance from being overloaded. This makes the entire web architecture more scalable and adaptable to changing needs.

- Less maintenance – The website owners don’t have to worry about purchasing, maintenance, troubleshooting and upgrading the underlying infrastructure. This is taken care of by AWS. They focus on enhancing and driving innovation.

- Cost-effective – Hosting the website on AWS is much more cost-effective as there is no infrastructure and maintenance cost involved. Resources were consumed on a pay-as-go and subscription-based model.

4.3.3 KEY ACHIEVEMENT

In the process of understanding the AWS architecture, perofming the application migration and under taking the various online courses, I was able to complete my “AWS Certified Cloud Practitioner Certificate”.

CHAPTER 5

CONCLUSION

After understanding, observing, and implementing the various functionalities and features of Cloud Computing, it can be concluded that cloud computing is helping customers overcome traditional data centers' challenges. Traditional data centers had many challenges limited scale, flexibility, a huge number of expenditures, lack of visibility, and security gaps, to name a few.

On the other hand, cloud computing provides customers with an environment that is much more flexible, global, and scalable. Cloud computing has two fundamentals of the ‘as-a-service’ model and ‘pay-as-you-go’ model, which are major customer attractions. With these principles, customers can get IT resources as a service. They can consume these services in a subscription-based model. With the ‘pay-as-you-go’ principle, customers enjoy much more flexibility as they can now pay only for the resources they consume. These two principles have made cloud computing a very attractive alternative to customers, mainly because with these two principles, customers can save huge amounts of money. This money saved can be invested back into the business to drive innovation and growth. Hence, cloud computing is completely in-line with the business of the customers.

More and more companies are adopting cloud computing today due to these benefits. Multiple cloud vendors are available in the market today to provide these cloud services. However, each vendor provides comparable services, pricing, and global reach. Amazon Web Services (AWS) is the world’s most adopted cloud network. Millions of customers across the globe are now connected to the AWS cloud. The major reasons for AWS's popularity are its large variety of services and its very big global scale. AWS offers more than 200+ services and is spread across 30 regions around the globe. The migration of a test application to the AWS cloud helped to understand and project these powerful services provided by AWS. Consuming these services is straightforward. AWS has the biggest market share today in the cloud computing market. AWS is receiving strong competition from Microsoft Azure and Google Cloud Platform. However, AWS is still a popular choice among customers and is expected to grow much more rapidly in the future.

FUTURE WORK

This project focused on an in-depth analysis of cloud computing, AWS cloud, and migration of a test application to the cloud to showcase the benefits of the cloud. The future work on the project can focus on an in-depth analysis of the AWS cloud with one of the major cloud vendors like Microsoft Azure. An apple-to-apple comparison of the two biggest cloud vendors will help to understand the architecture and services more deeply.

Further, the application migration can also include migrating a dynamic database with a microservice-based architecture. The migration of the same application can be carried out on Microsoft Azure cloud to understand the functionalities, similarities, and differences between the two cloud providers.

REFERENCES

[i] *Reference: https://www.cisco.com/c/en_ca/solutions/data-center-virtualization/what-is-a-data-center.html*

[ii] *Reference: https://www.cisco.com/c/en_ca/solutions/data-center-virtualization/what-is-a-data-center.html*

[iii] *Reference: https://www.temok.com/blog/datacenter-or-data-center/*

[iv] *Reference: https://www.zdnet.com/article/what-is-cloud-computing-everything-you-need-to-know-about-the-cloud/*

[v] *Reference: https://en.wikipedia.org/wiki/Cloud_computing*

[vi] *Reference: https://en.wikipedia.org/wiki/Cloud_computing*

[vii] *Reference: https://aws.amazon.com/what-is-cloud-computing/*

[viii] *Reference: https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-cloud-computing*

[ix] *Reference: https://www.zdnet.com/article/what-is-cloud-computing-everything-you-need-to-know-about-the-cloud/*

[x] *Reference: https://www.hpe.com/us/en/what-is/on-premises-vs-cloud.html#:~:text=On%2Dprem%20and%20cloud%20computing,a%20third%2Dparty%20service%20provider*

[xi] *Reference: https://www.educba.com/on-premise-vs-cloud/*

[xii] *Reference: https://www.dataversity.net/how-the-cloud-has-evolved-over-the-past-10-years/*

[xiii] *Reference: https://www.dataversity.net/how-the-cloud-has-evolved-over-the-past-10-years/*

[xiv] *Reference: https://www.researchgate.net/figure/The-Evolution-of-Cloud-Computing_fig1_228991214*

[xv] *Reference: https://www.oreilly.com/pub/pr/3333*

[xvi] *Reference: https://www.imperva.com/learn/availability/fault-tolerance/*

[xvii] *Reference: https://www.techrepublic.com/article/report-84-of-it-pros-will-move-more-workloads-to-the-cloud-by-2019-for-cost-savings-flexibility/*

[xviii] *Reference: https://www.businesstechweekly.com/operational-efficiency/cloud-computing/economics-of-cloud-computing/*

[xix] *Reference: https://technologyadvice.com/blog/information-technology/4-ways-cloud-computing-can-save-money/*

[xx] *Reference: https://www.vmware.com/topics/glossary/content/cloud-scalability.html*

[xxi] *Reference: https://blogs.mulesoft.com/api-integration/cloud/what-is-scalability-in-cloud-computing/*

[xxii] *Reference - https://www.nojitter.com/what-being-agile-company-means-and-why-it-matters*

[xxiii] *Reference:https://www.google.com/url?sa=i&url=https%3A%2F%2Fdigitaltransformationtrends.com%2F2019%2F12%2F04%2Fbenefits-distributed-cloud%2F&psig=AOvVaw12cYCf-sPWHtAIylNfLb1u&ust=1677308323170000&source=images&cd=vfe&ved=0CBAQjRxqFwoTCJCui8jKrf0CFQAAAAAdAAAAABAE*

[xxiv] *Reference - https://aws.amazon.com/types-of-cloud-computing/*

[xxv] *Reference: https://www.redhat.com/en/topics/cloud-computing/iaas-vs-paas-vs-saas*

[xxvi] *Reference: https://www.routexp.com/2021/11/software-as-service-saas-model-cloud.html*

[xxvii] *Reference - https://www.fingent.com/blog/cloud-service-models-saas-iaas-paas-choose-the-right-one-for-your-business/*

[xxviii] *Reference: https://azure.microsoft.com/en-ca/resources/cloud-computing-dictionary/what-is-iaas/*

[xxix] *Reference: https://www.cloudflare.com/en-gb/learning/cloud/what-is-iaas/*

[xxx] *Reference: https://www.fingent.com/blog/cloud-service-models-saas-iaas-paas-choose-the-right-one-for-your-business/*

[xxxi] *Reference: https://www.fingent.com/blog/cloud-service-models-saas-iaas-paas-choose-the-right-one-for-your-business/*

[xxxii] *Reference: https://www.cloudflare.com/en-gb/learning/serverless/glossary/platform-as-a-service-paas/*

[xxxiii] *Reference: https://www.fingent.com/blog/cloud-service-models-saas-iaas-paas-choose-the-right-one-for-your-business/*

[xxxiv] *Reference: The Basics of Cloud Computing Understanding the Fundamentals of Cloud Computing in Theory and Practice*

[xxxv] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xxxvi] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xxxvii] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xxxviii] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xxxix] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xl] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xli] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xlii] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xliii] *Reference: https://www.bmc.com/blogs/public-private-hybrid-cloud/*

[xliv] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xlv] *Reference: https://www.mongodb.com/cloud-database/private-public-hybrid-cloud*

[xlvi] *Reference: https://www.knowledgehut.com/blog/cloud-computing/top-cloud-computing-companies*

[xlvii] *Reference: https://www.knowledgehut.com/blog/cloud-computing/top-cloud-computing-companies*

[xlviii] *Reference: https://www.knowledgehut.com/blog/cloud-computing/top-cloud-computing-companies*

[xlix] *Reference: https://www.knowledgehut.com/blog/cloud-computing/top-cloud-computing-companies*

[l] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[li] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[lii] *Reference: https://aws.amazon.com/what-is-aws/*

[liii] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[liv] *Reference:*https://aws.amazon.com/what-is-aws/

[lv] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[lvi] *Reference: https://bigsteptech.com/blog/aws-stats-in-2023/*

[lvii] *Reference: https://www.simplilearn.com/aws-stats-article#:~:text=AWS%20Total%20Users,small%20and%20medium%2Dsized%20businesses.*

[lviii] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[lix] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[lx] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[lxi] *Reference: https://www.infoclutch.com/installed-base/cloud-computing-software/amazon-aws/*

[lxii] *Reference: https://en.wikipedia.org/wiki/Amazon_(company)*

[lxiii] *Reference: https://en.wikipedia.org/wiki/Amazon_(company)*

[lxiv] *Reference: https://en.wikipedia.org/wiki/Amazon_(company)*

[lxv] *Reference: https://techcrunch.com/2016/07/02/andy-jassys-brief-history-of-the-genesis-of-aws/*

[lxvi] *Reference: https://en.wikipedia.org/wiki/Amazon_(company)*

[lxvii] *Reference: https://en.wikipedia.org/wiki/Amazon_Web_Services*

[lxviii] *Reference: https://aws.amazon.com/about-aws/global-infrastructure/*

[lxix] *Reference: https://aws.amazon.com/about-aws/global-infrastructure/*

[lxx] *Reference: https://aws.amazon.com/about-aws/global-infrastructure/regions_az/*

[lxxi] *Reference: https://explore.skillbuilder.aws/learn/course/134/play/62437/aws-cloud-practitioner-essentials*

[lxxii] *Reference: https://aws.amazon.com/about-aws/global-infrastructure/regions_az/*

[lxxiii] *Reference: https://explore.skillbuilder.aws/learn/course/134/play/62437/aws-cloud-practitioner-essentials*

[lxxiv] *Reference:https://explore.skillbuilder.aws/learn/course/134/play/62437/aws-cloud-practitioner-essentials*

[lxxv] *Reference: https://aws.amazon.com/ec2/instance-types/*

[lxxvi] *Reference: https://aws.amazon.com/ec2/pricing/*

[lxxvii] *Reference: https://explore.skillbuilder.aws/learn/course/134/play/62437/aws-cloud-practitioner-essentials*

[lxxviii] *Reference: https://explore.skillbuilder.aws/learn/course/134/play/62437/aws-cloud-practitioner-essentials*

lxxix Reference: https://aws.amazon.com/products/networking/

lxxxReference: https://aviatrix.com/learn-center/cloud-security/aws-networking-service-offers/#:~:text=AWS%20networking%20services%20allow%20customers,to%20personal%20private%20virtual%20networks.

lxxxi Reference: https://aws.amazon.com/vpc/

lxxxii Reference: https://aws.amazon.com/vpc/

lxxxiii Reference: https://docs.aws.amazon.com/vpc/latest/userguide/configure-subnets.html

lxxxiv Reference: https://explore.skillbuilder.aws/learn/course/134/play/62437/aws-cloud-practitioner-essentials

[lxxxv] *Reference: https://aws.amazon.com/products/management-and-governance/*

lxxxvi Reference: https://aws.amazon.com/cloudwatch/

lxxxvii Reference: https://aws.amazon.com/cloudwatch/

lxxxviii Reference: https://aws.amazon.com/cloudwatch/

lxxxix Reference: https://aws.amazon.com/cloudtrail/

xc Reference: https://aws.amazon.com/cloudtrail/

xci Reference: https://aws.amazon.com/premiumsupport/technology/trusted-advisor/

xcii Reference: https://aws.amazon.com/premiumsupport/technology/trusted-advisor/

xciii Reference: https://www.jeffersonfrank.com/insights/top-reasons-cloud-migrations-fail

xciv Reference:https://veritisgroup.medium.com/top-8-significant-strategies-for-cloud-migration-considerations-aa0d98845981#:~:text=However%2C%20cloud%20migration%20is%20a,on%20the%20cloud%20migration%20initiative.

xcv Reference:https://aws.amazon.com/blogs/enterprise-strategy/6-strategies-for-migrating-applications-to-the-cloud/

xcvi Reference:https://aws.amazon.com/blogs/enterprise-strategy/6-strategies-for-migrating-applications-to-the-cloud/

xcvii Reference: https://aws.amazon.com/blogs/enterprise-strategy/6-strategies-for-migrating-applications-to-the-cloud/

[xcviii] *Reference:https://wire19.com/amazon-microsoft-and-google-cloud-infrastructure-market/*

[xcix] *Reference:https://www.bmc.com/blogs/cloud-availability-regions-zones/*

[c] *Reference:https://www.bmc.com/blogs/aws-vs-azure-vs-google-cloud-platforms/*

[ci] *Reference:https://www.bmc.com/blogs/aws-vs-azure-vs-google-cloud-platforms/*

[cii] *Reference:https://www.bmc.com/blogs/aws-vs-azure-vs-google-cloud-platforms/*

[ciiiciii] *Reference: https://www.bmc.com/blogs/aws-vs-azure-vs-google-cloud-platforms/*

[civ] *Reference: https://kinsta.com/blog/wordpress-statistics/*

[cv] *Reference: https://kinsta.com/blog/wordpress-statistics/*

[cvi] *Reference: https://aws.amazon.com/what-is/lamp-stack/*

[cvii] *Reference: https://phoenixnap.com/kb/what-is-a-lamp-stack*

[cviii] *Reference: https://www.serverwatch.com/guides/lamp-stack/*

[cix] *Reference: https://aws.amazon.com/what-is/lamp-stack/*

[cx] *Reference: https://aws.amazon.com/what-is/lamp-stack/*

[cxi] *Reference: https://aws.amazon.com/what-is/lamp-stack/*

[cxii] *Reference: https://aws.amazon.com/what-is/lamp-stack/*

[cxiii] *Reference: https://aws.amazon.com/what-is/lamp-stack/*