ISSN 1533-9211

DOI 10.17605/OSF.IO/7E5Q9

INFLUNCES OF GPU CHALLENGES ON CLOUD COMPUTING WITH WEB TECHNOLOGY APPLICATIONS

SHILAN AHMED MOHAMMED

ITM Department, Technical College of Administration, Duhok Polytechnic University, Duhok, Iraq.
Email: shilan.itm@gmail.com

ZAINAB SALIH AGEED

Computer Science Department, College of Science, Nawroz University, Duhok, Iraq.
Email: zainab.ageed@nawroz.edu.krd

LOZAN M. ABDULRAHMAN

ITM Department, Technical College of Administration, Duhok Polytechnic University, Duhok, Iraq.
Email: lozan.abdulrahman@dpu.edu.krd

TEBA MOHAMMED GHAZI SAMI

Computer Science Department, Faculty of Science, University of Zakho, Duhok, Iraq.
Email: teba.sami@uoz.edu.krd

MOHAMMED JASIM AHMED

Erbil, Iraq. Email: mja.ise.1981@gmail.com

Abstract

There have been a growing number of academics and developers who have appeared in recent years, and they have spent a significant amount of their time and attention to researching GPU, as well as the numerous sectors in which it plays an essential and successful function. This is in part due to the immense progress that the world is seeing on a daily basis, which includes a variety of disciplines of study such as medical, engineering, space, and precise sciences that demand high-resolution visual representation and vast graphic processing power. This article acts as a review and provides a condensed description of the development, applications, problems, and related research into GPU across a wide number of various industries. It should not come as a surprise, given all of these benefits, that cloud computing has become the standard practice in its business in such a short length of time. The movement of corporate operations to the cloud, which is being done by a growing number of companies, is being done in an effort to minimize the amount of time spent on infrastructure maintenance. This migration is taking place for the following reasons: As a consequence of this, preserving a cloud environment in such a manner that it continues to operate normally is a tremendously challenging endeavor. If you want to increase cloud performance and minimize the amount of administrative stress you suffer, you need a reliable cloud monitoring solution. The cloud monitoring service is an investment that is well worth making since there is a possibility that it will boost performance and minimize the amount of time that is spent managing. One of the most important roles of cloud monitoring is the management of Quality of Service (QoS) metrics for cloud-hosted, virtualized, and physical services and applications.

Keywords: GAS model, Graph processing, GPU, Graph, BSP model

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

1. INTRODUCTION

A graphics processing unit, sometimes known as a GPU, is a kind of computer chip that is intended to do mathematical operations at a high speed in order to generate visuals and images. Graphics processing units, sometimes known as GPUs, are increasingly being used in computers for both professional and recreational purposes. GPUs have typically been responsible for the rendering of two-dimensional and three-dimensional graphics, as well as animations and video. The graphics processing unit, or GPU, is a specialized electrical circuit that is used in embedded devices, mobile phones, social networks, and the analysis of web links. Its purpose is to rapidly manipulate and change memory in order to assist in the synthesis of images that are to be transmitted to a display. Today's graphics processing units (GPUs) are very effective at manipulating computer visuals, and their highly parallel structure makes them more powerful than general-purpose CPUs for algorithms that process huge blocks of data in parallel. GPUs are used in desktop PCs, laptops, and gaming consoles. A general-purpose graphics processing unit (GPU) is being used in lieu of the stream processor. This GPU is also acting as the "left behind" stream processor. In a manner similar to that of a conventional personal computer, the NVIDIA computational architecture may serve a multitude of functions. When interacting with modern graphics accelerators in this manner, it is feasible to make use of these accelerators for computation that is unrelated to graphics [1]. Because of NVIDIA's CUDA, a considerable portion of a standard C program may now make advantage of the stream processors located on the GPU. The creation of new, creative applications for specialized hardware is made possible by this technology. As a consequence of this, a standard C program has the ability to execute in parallel on many GPUs, which enables it to rapidly analyze large matrices without the need to recompile them. It can even switch between utilizing the CPU and using the GPU depending on the requirements of the task at hand. Applications that are based on the CPU have the option of utilizing the CUDA (Compute Unified Device Architecture) API rather than employing a graphics application programming interface (API). This will allow the CPU-based applications to gain access to the GPU's capabilities for more general computation [2].

A mathematical structure known as a graph is made up of a collection of nodes, also known as "vertices," and the lines, also known as "edges," that connect those nodes to one another. In more recent times, the concept of using such graphs as linkages between objects has been applied to real-world challenges.

The following are some examples of real-world applications for using graphs with GPUs:

- Electronic Design Automation
- Bioinformatics
- Data Science, Analytics, and Databases
- Media and Entertainment
- Computational Fluid Dynamics

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

- Machine Learning
- Computational Finance
- Defense and Intelligence
- Imaging and Computer Visions
- Medical Imaging

In addition, graphs are commonly used by computer systems and are used to represent diverse structures. Also, graphs are frequently used in compilation optimization [3].

2. BACKGROUND THEORY

This is directly pushing the information technology industry to lean more toward the usage of remote computing, which is a direct effect of this. Consequently, this is a direct outcome of this. In today's modern workplace, it is normal practice for workers to make advantage of remote access to office workspaces in order to get their actual work done. This is because remote access to office workspaces allows employees to work more efficiently. This is due to the increased prevalence of technology advancements in today's workplaces. Word processors, spreadsheets, and several other forms of productivity software are likely to be present in the majority of these businesses. However, it is still challenging to set up a multi-user environment for effective remote work that takes use of applications that are visually demanding (CAD/CAM and GPU computing, for example). The reason for this is due to the fact that these applications need a considerable amount of graphics processing power. However, these cloud-based services come at a premium price and are only compatible with a limited number of the available pieces of hardware, operating systems, and directories. There are a number of companies that provide cloud-based services to virtualize graphics processing units (GPUs). These services are provided by a number of different companies. Even though market-based solutions are often required to solve problems, this does not necessarily make them the optimal choice [4].

Over the course of the last several years, a variety of approaches to remote computing have emerged as successes at the cutting edge of information technology in both the business and academic spheres. Quite a few distinct businesses are responsible for the development of these methodologies. These kind of technical developments have resulted in the birth of a great deal of new terminology, some examples of which include "cloud computing," "software as a service," and "platform as a service," amongst a great deal of other phrases. In spite of the fact that both network-transparent graphics (X Window System) and remote access (telnet) were invented more than three decades ago, both technologies continue to advance in order to fulfill the requirements of contemporary life [5].

When the client browser has an adequate degree of compatibility with HTML5, it is possible to use word processors, spreadsheet applications, and other products that are analogous in nature in an efficient manner when they are accessed remotely through the utilization of a web interface. This is made possible when the client browser has an adequate degree of HTML5

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

compatibility. This is true even if the browser being used by the client does not have compatibility with HTML5. On the other hand, since many applications are very visually intensive, graphics processing units (GPUs) are required for either interactive rendering or computation that is based on GPUs. This is the case because GPUs are required to handle the visual demands of many applications. This is because graphics processing units are the foundation upon which GPUs are built. Even the most pessimistic predictions demonstrate that a connection capable of Gigabit Ethernet can send an interactive video stream with a resolution of 1024 by 768 pixels and a color depth of 16 bits per pixel. This is the case even when the estimations are the most optimistic they can be. Even if the estimations are pushed to their limits, this conclusion still holds true [6].

If an advanced video compression technology is used, the amount of bandwidth that is required for the transmission of a video may be decreased by a factor of 10 or more. This reduction in bandwidth requirements is possible. Because of this, we would be able to utilize the available bandwidth in a manner that was far more effective. Having said that, it is easy to jump to the conclusion that it is easy to create a graphical environment in which multiple users are able to interact with one another, despite the fact that this is in no way at all the case in reality. This is because it is simple to create a graphical environment in which multiple users are able to interact with one another. The creation of a graphical environment in which several users are able to communicate with one another is not a straightforward process; this is a fact. The primary reason for this is the complexity that is involved in the process of setting up an interactive multi-user graphical environment, which includes but is not limited to elements such as the hardware layer (GPU drivers), the windowing (and, sometimes, audio) system, and the network layer; the input events from the client must be delivered undistorted (a real problem with the Virtual Network Computing - VNC -protocol), and they must be processed with minimal delay. Another reason for this is that interactive multi-user graphical environments Moreover, this is due to the fact that interactive multi-user graphical environments are becoming increasingly common. The method of streaming video is far less difficult than the process of developing an interactive graphical environment, which is a technique that requires much more effort [7].

There are now just three commercial solutions available for use with this objective; the corporations Microsoft, Citrix, and VMware are responsible for developing these solutions. All of these approaches need expensive graphics processing units (GPUs), in addition to specialized operating systems (OS) that are responsible for providing directory services. Strong integration with proprietary directory services may be beneficial in business organizations; however, carrying out such an activity in the context of a university may only result in the addition of extra administrative responsibilities (though Linux and FreeBSD clients can nicely coexist with Microsoft Active Directory), It is normal practice to use the words "virtual workplace" and "remote workplace" interchangeably, and the fact that this happens is not a secret. This is a behavior that has become commonplace. However, the basic issue is that client-server solutions for graphical remote desktop environments are notoriously difficult to create. This is the source of the difficulty. According to its "definition" on Wikipedia as of July 2015, the former refers to "a workplace that is not located in any one physical region." This is only

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

one example among many. This phrase does not, by itself, communicate a very substantial amount of the division of labor that occurs in an environment that is networked. The system that is not linked with a local console will henceforth be referred to as a "virtual workplace" or a "remote workplace," respectively, to characterize the situation in which essential components of relevant tasks are done on a regular basis. These terms will be used interchangeably in the future. In the future, we are going to use both of these words interchangeably. This adjustment will take effect as soon as possible. A text console or graphical terminal session that is initiated remotely to a network server for administrative reasons is an example of something that does not properly qualify as a virtual workplace. This kind of session is often used for administrative purposes. This is as a result of the fact that a system that has been adequately prepared should not have too many sessions of this sort opened at the same time. The text interface of the packet batch scheduler of the HPC cluster, on the other hand, exemplifies a remote office in its most unadulterated form. The recommended meaning of a virtual workplace cannot be entirely defined unless the ambiguity associated with the word "important" is also eliminated. Until this ambiguity is addressed, the definition cannot be considered complete. This is due to the fact that the meaning of the term "important" may refer to more than one concept at the same time. When that time comes, the suggested interpretation will finally be able to be labeled in the way that is most suitable [8].

It is not entirely correct to say that not all sorts of workloads or applications are able to make advantage of GPU instances. Be sure that you don't lose sight of the fact that these instance types are, in general, the better choice for activities that need a lot of computational power before you start delving deeper.

The capability of graphics processing units, often known as GPUs, to carry out parallel processing makes them an excellent choice for some applications that are concerned with business analytics. There is also the possibility that some applications pertaining to artificial intelligence might gain an advantage from the employment of GPUs. GPUs may also be beneficial for other applications, such as those used for the development of films, virtual desktop infrastructure, and engineering simulation, which all need a significant amount of processing power. GPUs may be used for several kinds of applications. Additionally, companies that depend on supercomputing for academic research may also discover that GPUs are advantageous to their operations. This is because GPUs are able to process data much more quickly [9].

On the other hand, cloud instances that are based on GPUs tend to be more expensive than their virtual counterparts that are based on CPUs. This is because GPUs are more complex than CPUs. Even if it is possible that the expenses of adopting the technology within the business itself will be reduced, this will not change the fact that this is the case. It is probable that certain IT departments may have to go through a learning curve in order to take advantage of this, since GPUs are capable of handling a wide variety of various sorts of workloads [10].

There is a significant pricing gap between AWS, Azure, and Google for the usage of GPUs, with AWS's rates ranging from $0.70 to $0.90 per GPU per hour and Azure's and Google's prices ranging from $0.50 to $0.80 per GPU per hour. The costs for GPU use at AWS are much

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

higher. On the other hand, the starting price for one of Amazon Web Services' virtual CPU-based general-purpose instances is merely $0.0058 per hour.

In addition, applications for high-performance computing (HPC), which frequently make use of graphics processing units (GPUs), require personalized programming in addition to expertise working with specialized tools and frameworks such as Apache Spark, TensorFlow, and Torch. GPUs are referred to as graphical processing units. In order for companies to boost the possibility that their staff will be able to efficiently create and manage a variety of applications, it may be essential for the firms to shell out money for employee training and certification [11].

The most well-known companies that provide cloud computing services give each of their customer's access to their very own unique instance of GPU hardware. AWS makes its Elastic Graphics Processing Units (EGPUs) accessible to customers, and these EGPUs have the capability of establishing connections to EC2 instances that are running in any location within the AWS cloud. Users need to establish the compute, memory, and storage needs that are presented by the workload before they can decide which sort of EC2 instance type to utilize for their application. EC2 instance types include: In order to proceed with the option, this is a prerequisite that must be met. A multitude of different GPU memory capacities are available to be used via AWS, beginning at 1,024 MB and going all the way up to 8,192 MB.

Customers have access to a variety of virtual machines, including those in the N-series. Two examples of these virtual machines are the NC and NV GPU instances that Azure offers. Applications that place a high demand on both computing and networking are a perfect match for the NC sizes, which are powered by Nvidia's Tesla K80 card. These applications are a good fit for the NC sizes. To put it another way, the NC sizes are well suited for applications such as these. NV sizing, which is used for visualization, gaming, and encoding, is able to be done with the assistance of Nvidia's Tesla M60 GPU card and Nvidia Grid. Customers will not be able to utilize GPU instances unless they have first constructed such instances using Azure Resource Manager. Until then, customers will not be able to use GPU instances. Customers will not have access to such instances until such time as it is determined [12].

The Google Compute Engine platform has features that are analogous to those that have been outlined in this piece of writing. It is possible to utilize Nvidia Tesla P100 and K80 GPUs in non-shared-core preset instances as well as in custom instances. This functionality is also available for instances with shared cores. This functionality is offered for use with bespoke instances. Users have a duty to be aware of the numerous limitations that are currently in effect, such as the fact that GPU instances may have a maximum amount of RAM that does not exceed 208 gigabytes. It is the users' obligation to be aware of these limitations [13].

Application management and development with a focus on cloud-based platforms: When designing, deploying, and maintaining cloud applications, it is vital to pay particular attention to how much of the cloud's resources are consumed as well as how well the apps themselves are developed in order to fulfill performance and reliability standards. This is important in order to verify that the apps will perform as intended in the manner that is anticipated. In light of these circumstances, the IT department will need to devise a strategy for the production of

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

cloud-based applications as well as a plan for the management of these applications, and they will need to continue to maintain this plan. In addition, they will need to design a plan for the management of these applications [14].

"Enterprise computing" refers to the practice of utilizing computers within significant organizations for the purpose of carrying out data processing. This practice is also known by the terms "information management" (IS) and "information technology" (IT), both of which refer to different aspects of the same field (IT). In the 1960s, not long after the introduction of mainframe computers, businesses started making broad use of computers for the first time in order to process data. These computers were mostly used in corporate settings. This particular application of computers took place inside the realm of enterprises. The broad availability of new technological capabilities has had a considerable impact on the paradigm of business programming, which has been dramatically transformed as a consequence. The paradigm has been significantly disrupted as a result of this development. Large mainframe computers have been mostly rendered obsolete by client-server architectures ever since the widespread availability of personal computers in the 1980s. This transformation happened not long after the first personal computers were introduced onto the market. In the 2000s, user-centric web-based commercial systems and e-commerce websites began to supersede client-server architecture as the major kind of commercial software architecture. Client-server architecture had been the prevalent type of commercial software architecture since the 1990s. Within the framework of commercial software, this change took place. During the decade of the 1990s, when this event occurred, people were just starting to become more acquainted with the internet. This was a time when people were just beginning to become more familiar with the internet [15].

As a direct consequence of these transformations, corporate networks are now bigger and more comprehensive than they have ever been, and they have a greater capacity to give internet access to their staff members than they ever had before. While this is going on, the complexity of these responsibilities, as well as the associated expense, has increased in the following ways: Individuals and companies in every region of the world put an incredible amount of money into the area of information technology each and every year. This category encompasses a wide range of actions, including the acquisition of new hardware and software as well as an increase in the number of devices that are connected to the network. Other examples of actions included in this category include expanding the physical footprint of the business. In addition to that, this category also includes a rise in the number of workers who perform their duties outside the office (in-house or out- sourced). Cloud computing, if it is effective, has the potential to dramatically revolutionize commercial computing by turning it into a service that can be accessed over the internet rather than a physical piece of equipment. This would be a significant departure from the traditional model of commercial computing. If cloud computing were to become widespread, this would be the situation. In the event that anything similar does occur, the business sector of the computer industry will never be the same again [16].

Along with the internet-based delivery of prepackaged computing services, a pricing model in which consumers pay only for the cloud resources that they actually use, and the capability for

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

end-users to instantly supply the resources that they demand, the capability to instantly supply the resources that end-users demand is one of the most important aspects of cloud computing. The provision of instantaneous resources to end users is an additional feature of cloud computing, which enables users to fulfill their own resource requirements. As a result of the pooling of their computing resources, cloud services may be able to realize significant economies of scale. This is something that remains to be seen, but it is a possibility. After then, it's possible that these cloud services will be able to transfer these savings to commercial information technology. When individuals talk about their excitement for cloud computing right now, it is not surprising that the possibility of cost reductions is one of the primary things that comes to mind as a significant driving factor in their minds as one of the reasons why. This is due to the fact that computing in the cloud has the potential to save expenses. Cloud services have made it feasible to mine data at scale that were not before reachable by combining massive quantities of processing power in centralized pools. This was previously only possible via the use of traditional data mining techniques. Data mining at sizes that were previously inaccessible has become feasible as a result of this development. This strategy led to the creation of a wide variety of unique programming models and development methodologies, many of which are now grouped together under the umbrella term "cloud computing." These technological developments make it possible to carry out calculations on a huge size and significantly enhance the efficiency of the process by which new software is created. Additionally, these advancements make it possible to conduct out computations on a massive scale. In addition, as a result of these developments, it is now feasible to carry out calculations in real time [17].

As a direct result of the meteoric rise in popularity of cloud computing over the course of the previous few years, many large multinational organizations are now engaged in the process of fundamentally reevaluating their strategy for information technology. This is happening as a direct consequence of the meteoric rise in popularity of cloud computing. As a result of the extraordinary advantages offered by the "as a service" paradigm in terms of investment, distribution time, and scalability, new (mobile) solutions have been widely disseminated, and emerging advancements such as big data, the internet of things (IoT), and machine learning (ML) have been brought to the forefront of attention. The "as a service" model is largely responsible for bringing these recent innovations to the forefront of public attention [18].

These developments are indications that the available resources may be enhanced on a global basis. This is a positive sign for the future of the cloud computing industry and increases the number of opportunities that are available to specialized companies that conduct their operations on a worldwide basis [19].

In order to give users the best experience possible and give customers the best experience possible, new computer and networking technologies are currently being developed to make use of infrastructure that is shared on a global scale. This is being done in order to provide users with the greatest experience possible. Because of the fact that the delay has been decreased, the immediate and inescapable result is that: On the other hand, assessing the quality of end-user services in a system that is composed of many clouds presents a number of significant

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

challenges (QoS). As a direct result of its growing significance within the industry, larger suppliers are becoming more dependent on SLM in order to properly carry out their business operations. This is a trend that is expected to continue. To a far greater extent than it is influenced by internal organizational or technical factors, the acceptance of cloud-based services among small and medium-sized businesses (SMEs) is mostly determined by external circumstances. This case demonstrates the significance of environmental effects in the process of small and medium-sized firms (SMEs) in India adopting new technology better than any other, which is why it is the finest example of its type and the largest example of its kind overall. In conclusion, this is the way that is the most successful for highlighting the importance of the effects that the process of small and medium-sized firms (SMEs) in India embracing new technology has on the environment. Focuses on the value of the data as well as an analysis of the load capacity of the cloud storage, The findings of the computer research shed light on the inner workings of the "cooling" process that was being investigated. It has been determined how some QS models that make use of queues may be modified to be more relevant to the world as it really exists. A number of calculations are performed in order to identify the bare minimum standard of service that must be provided in order to meet the predefined level of success. The results of these calculations are used to establish the bare minimum standard of service that must be provided. The choice of a data center to offer computational capacity is determined by the data center's available power, and the great majority of companies that provide services related to data centers advertise that their prices are flexible depending on the requirements of the customer. There is a significant number of data centers located in various parts of the world [20].

Many companies, despite the fact that enterprise cloud infrastructure has significant benefits in terms of productivity, efficiency, and cost savings, continue to struggle with significant obstacles when attempting to embrace the cloud. This is despite the fact that enterprise cloud infrastructure has these benefits. The fact that business cloud infrastructure offers these advantages does not change the reality that this is the case. The proliferation of social video and streaming media services has resulted in an appreciable growth in the degree of expertise that can be found in business cloud computing. Because users have come to expect nothing less than a flawless experience across all digital mediums, network congestion poses a serious threat to the consistency of consumer-facing digital services such as websites, mobile applications, and digital video. This is because users have come to expect nothing less than a flawless experience across all digital mediums. This is due to the fact that consumers have become used to expecting nothing less than a perfect experience across all forms of digital media. Users have also come to expect nothing less than a flawless experience across all digital mediums. This is what they have come to expect. This is something that users have learned to anticipate. Due to the prevalence of hazards such as distributed denial of service attacks and other complex dangers, companies who utilize commercial cloud storage services are exposed to a significant risk to the completeness and accuracy of the data they store. This risk may be considered high [21].

Businesses who are interested in lowering the hurdles that prevent them from entering the commercial cloud computing market may find it helpful to make use of the monitoring tool

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

technologies that are now at their disposal. Because it is the biggest and most comprehensive cloud-based infrastructure in the world, enterprise cloud computing is made more scalable, efficient, and safe as a consequence of the availability of this cloud-based infrastructure. This infrastructure is put to use in the process of providing and speeding a number of different technologies that are associated with corporate cloud computing. This is because the infrastructure is used not just to handle data but also to store it in the cloud, where it can be accessed for processing. Because of this, companies now have the capability to extend their networks into the cloud and to control traffic through the Internet. This provides businesses with the ability to guarantee that their workers, business partners, and customers have protected access to the services that the firm regards to be of the utmost significance.

With the assistance of web performance solutions, the speed at which websites and other types of online applications function may be improved. Because of this, businesses are now able to provide their consumers access to their services from any device, without sacrificing speed, reliability, or security. When businesses use media distribution technologies such as a live streaming server, they are able to provide their customers with a viewing experience that is seamless, provide quick access to high-definition multimedia content on any device, and have the scalability to meet peak demand for their products and services. In addition, the businesses are able to meet the demands placed upon them by their customer base [22].

The use of cloud computing provides websites and data centers with protection against the loss of data in addition to uptime. This protection may be acquired without the need of installing any extra software or hardware. By incorporating cloud networking technology into the modernization of its corporate network, an organization has the opportunity to improve application performance, save operating expenditures, and expand cloud accessibility. Advances on the part of service providers that make it easier to establish and administer networks for the aim of supplying a variety of services to clients in order to meet their diverse needs. It is imperative that all major businesses regularly evaluate the viability of their existing information technology (IT) infrastructures in order to ensure that critical business operations continue to run smoothly. This is because the client-server model is quickly being replaced by the internet of things and cloud computing. In addition to this, it is of the utmost importance that all of the main companies guarantee that their most important business processes continue to function without a hitch. This is because the landscape of information technology is quickly shifting away from client-server architecture and toward internet of things and cloud computing. The reason for this is due to the fact that cloud computing and internet of things are both forms of distributed computing. It is of the utmost importance to make certain that all pre-existing business processes continue to function faultlessly even after new technologies have been implemented [23].

Because of the need to give help in a wide variety of situations, including quickly expanding technological environments, the "enterprise" part of information technology was absolutely necessary.

Have benefited from the building of an overall framework for how the company's software satisfies business needs, how it is applied structurally, and how data is genuinely exchanged

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

across the different business applications. This structure was conceived and designed by the firm. Regulation and standardization of the availability of newly established technological platforms are very required in order to guarantee that these platforms may be used. This is necessary to ensure that these platforms may be used. When conducting research into the potential applications of cloud computing by large companies, it is essential to examine the transition in infrastructure model from the point of view of the numerous business operations that will be affected by the change. This is because the change will affect many of the business operations that are currently being carried out. This is because the shift will have an effect on the way those particular company processes are carried out [24].

The utilization of computing that is carried out through the cloud offers a great deal of space for advancement. The years that have passed since the introduction of cloud computing have shed light on the shortcomings of the conventional IT infrastructure. Because their technical environments are not well-equipped to recognize these shifts and put acceptable solutions into action, businesses are finding it increasingly difficult to adapt to changes in the market as well as improvements in technology. This is because the market is constantly shifting, and so are technological advancements. As a consequence of this, maintaining a competitive advantage will require a great deal more effort than before. Because of this, it is far more difficult for businesses to adjust when they are confronted with altering conditions. Computing in the cloud makes it possible to develop an information technology infrastructure that is both adaptable and efficient; this architecture was created with one objective in mind: to assist organizations in reaching their potential for success and expansion [25].

Interoperability is improved when people who belong to a variety of organizations and teams are able to access and make use of the same data that is stored in the cloud. This adds to an overall improvement in the state of interoperability. As a direct consequence of this fact, individuals will have an easier time working together toward a common goal. Because of this strategy, remote workers are able to gather data in a timelier manner, and their capacity to connect with management has been much improved. This is performed by lowering the amount of lag time that is often associated with conventional IT models, which in turn helps to speed up the process. This is done in order to achieve the aforementioned goal. It may be more difficult for staff employees working at a distant area to carry out their responsibilities in an effective way as a result of these delays. It makes the processes easier to understand, which in turn makes it possible to create more output in a shorter amount of time [26].

• The capacity to continue functioning well while simultaneously expanding in size: The specifications that are needed in the industry are always shifting. As a direct consequence of this fact, the architecture of the cloud was developed so that it is capable of expanding whenever the need for such capabilities arises. When a firm expands, there is a corresponding increase in the demand for a bigger volume of data storage as well as an expanded capacity for the underlying network. Both of these requirements must be met simultaneously. With auto-scaling, businesses are able to adjust the amount of cloud resources they use to effectively meet the requirements of their operations, regardless of whether this means increasing or decreasing the amount of cloud resources they utilize. This flexibility allows businesses to effectively meet

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

the demands placed on them by their operations. When a website's architecture is built using cloud computing, load times for the website are sped up, and the amount of downtime experienced is reduced to an absolute minimum [27].

• Software that is capable of doing frequent automatic software updates: To ensure that the requirements of businesses are met, a huge number of cloud service providers are responsible for maintaining the functioning of the service. The maintenance of cloud servers is an ongoing process, which means that the resources that would have been required to accomplish the activity on-premises are no longer required [28].

In addition, companies that have made a commitment to the practice of corporate social responsibility and are concerned with reducing their carbon footprint will realize that cloud-based solutions are an excellent option for their organization and should give serious consideration to making the switch. Businesses that have made a commitment to practicing corporate social responsibility and are concerned with lowering their carbon footprint. The broad use of cloud storage has the potential to cut the number of essential components of information technology infrastructure in the workplace by a substantial margin, and in certain cases it may even do away with the need for these parts entirely. It is possible that the fact that cloud storage may diminish or perhaps eliminate the need for physical storage devices was one of the elements that led to this problem. If this is the case, then it is very plausible that this was one of the factors. They are only required to hand over money to cloud service providers in line with the "pay as you go" pricing model that these cloud service provider's use. In other words, companies are only required to pay for the cloud services that they actually use. This indicates that companies are only obligated to send over money when they are really using the cloud service that they have subscribed to. It may make sense for a company to take advantage of the scalability that is provided by its storage system when the natural growth and development of a business inevitably results in the requirement for extra capacity. This is the case when the need for additional capacity cannot be avoided. The capacity of a system to grow or shrink in size in response to changes in requirements is referred to as its scalability. Now, so long as they are connected to the network, workers may do their tasks from wherever they happen to be, regardless of where they are physically located. This is now a choice that can be made and should be thought about as a result of the fact that information that is kept in the cloud is kept at a place that is remote on the internet [29].

The aggregate name for the many different online computer resources that are made available to customers of a variety of businesses via the use of the internet is known as the "digital cloud." Several different types of businesses make these materials accessible to the public. When it comes to computing in the cloud, public clouds have an advantage over private clouds since anybody who wants to use them or pay for them is allowed to do so. Private clouds, on the other hand, only allow users who have been invited by the cloud provider. On the other hand, private clouds are off limits to everyone who wants to utilize them. On the other hand, the only people who are permitted to utilize private clouds are those who have paid to get access to such clouds. The user may have access to these programs at no cost, or they may be bought for a charge, the amount of which is decided by the total quantity of time, space, or data that is

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

delivered. Alternatively, the user may have access to these applications at no cost at all. The removal of the need for companies to acquire, maintain, and keep up to date the hardware that is stored on their own premises is one of the most significant advantages that cloud computing gives to businesses. This eliminates one of the most significant costs that businesses face. Businesses will have an easier time keeping up with the ever-increasing demands that are being put on them by their data as a result of the fact that their storage capacity can be expanded and that their RAM can be upgraded [30].

The hybrid cloud is a combination of public and private cloud components. It utilizes just a subset of the hardware, software, and network resources that are made available by cloud computing. Because of the "cloud voodoo" notion, there is the potential for both public and private clouds to transform into a variety of various shapes in order to satisfy varying needs for the amount of computing power. Because this traffic isn't coming from the public cloud, companies will need to increase the capacity of their networks in order to keep up with the increased demand for cloud storage and on-premises storage. This is necessary in order for enterprises to be able to meet customer expectations. This is essential for companies to do in order to live up to the standards set by their customers. In cloud computing infrastructures that are able to accommodate numerous tenants at once, services are often bought on an à la carte basis. You may be able to reap the advantages of cloud computing in many contexts without having to take the precautions necessary to protect your data from the security threats that are often connected with it. This is something that is becoming into a more widespread occurrence [27].

The phrase "infrastructure as a service," which is usually abbreviated as "IaaS," is used to refer to cloud-based server and storage capabilities, as well as virtualized network connectivity. This term is widely referred to. This notion, which is employed in computers, is referred to as "as a service," and this word is used within the context of that framework. These expression, which depicts the process of physically gaining something, is most often used in the phrase "getting your hands on this goods," which also contains the idiom. Even while the client is not directly responsible for the maintenance of the cloud resources, they do have some control over the servers, operating system, and apps that are given in connection with the service. In spite of the fact that the client is not directly responsible for the upkeep of the cloud resources, this is nonetheless the case [31].

An external service provider is responsible for assuming these commitments on behalf of the customer, but they are not responsible for the physical or logical infrastructure (including hardware, software, servers, and data storage, among other things) that is necessary to provide the service. The customer is not accountable for any of these duties or obligations in any way. There is not a single one of these aspects that may in any way be construed as being within the purview of the responsibility that is held by the consumer. The service provider is responsible for both the hosting of the program as well as the regular generation of backups of it. This responsibility falls within their jurisdiction since they are the ones providing the service. Amazon Web Services is the most well-known corporation that offers infrastructure as a service, which is also referred to as IaaS rather often. This service is also offered by several

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

other companies (AWS) [32].

The end user is relieved of the duties connected with the creation and management of the platform when cloud computing is carried out using the "platform as a service" (PaaS) model. This enables them to turn their attention to other matters. The term "platform as a service," which literally translates to "platform as a cloud service," is widely used to refer to configurations much like the one shown here. Instead, due to the fact that they are the ones who are in charge of delivering the service, the provider bears the whole of the responsibility for this situation. Developing software that is capable of working inside a network that is hosted in the cloud and offering great performance while doing so is the goal of this endeavor. Because customers are ultimately responsible for the amount of money that is spent on the services that they pay for, it is essential that they have the opportunity to provide feedback about the planning and execution of the services that they purchase [33].

Users that make use of software that is delivered in the form of a service (also known as SaaS) are able to get the maximum amount of benefit from the computer programs in which they have previously made a financial investment. Clients of software as a service are freed of the burden of having to worry about downloading updates or buying new versions of the program. This is because SaaS customers do not have to bother about these things. Because of this, it is now possible that this will happen. When it comes to software as a service (SaaS), users are not obliged to take on the administrative effort of setting up and maintaining their own physical infrastructure. As a result, users are able to get the most out of the software that they have already paid for. Users are able to get the most out of the program that they have already paid for by taking use of this feature. By using this feature, customers may get the most out of the software that they have already paid for, giving them more bang for their buck. They now have more extra time on their hands, which enables them to concentrate their attention to getting the most out of the program that they are using in order to maximize the advantages that they gain from it. As a result of this, businesses are now in a position where they are able to focus a greater portion of their attention on the production and marketing of new products, while simultaneously reducing the amount of focus that they pay to the routine maintenance of their systems and the design of new software. This is because businesses are now in a position where they are able to focus on the production and marketing of new products because of this. Both Amazon Web Services' (AWS) Elastic Beanstalk and Google's App Engine are well-known examples of popular platform as a service (PaaS) offerings. Both companies are owned and operated by Google (PaaS) [34].

"Software as a service," or "SaaS" for short, is an industry term that refers to the practice of providing customers with access to software by means of a subscription service. This kind of business model is known as "software as a service." This tactic is usually condensed into a shorter version. This model has been streamlined into a more condensed form for your convenience. Users who back up their information to the cloud don't need to worry about downloading or updating any software on their personal computers (PCs), since this is handled automatically. Users are free to retrieve any of their previously stored data anytime they choose. This is as a result of the fact that the cloud service takes care of each and every one of

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

those tasks on their behalf. They also don't have to put in the time and effort necessary to learn how to use any brand-new applications, which is yet another way in which they save both time and effort in the process of doing so. Because the software is stored on a different server, users have the choice of accessing it through an application programming interface (API) or a web browser. This is owing to the fact that the software is housed on a separate server (API) [35].

When the software as a service delivery model is used, the provider of the cloud storage is the one who is completely responsible for the encryption. In addition, the provider of the cloud storage is responsible for the hardware, middleware, and application services. "Software as a service," sometimes abbreviated as "SaaS," is a business model that gives firms the ability to access and make use of computer programs that are hosted on distant servers. The phrase "software as a service" (sometimes abbreviated as "SaaS") is often shortened. Because of this, not only will there be big enhancements in the effectiveness of administration, but also there will be large rises in the amount of joy that will be experienced by consumers as a direct result of this. " "Software as a service," sometimes known as "SaaS," is an abbreviation "refers to a group of software programs that includes, amongst others, Salesforce, Office 365 from Microsoft, and G Suite from Google [36].

Monitoring is an essential aspect of both the engineering and administration of today's increasingly complex networking infrastructures, which makes monitoring an important part of both of these professions. As a result, monitoring is an essential part of both of these fields. This is especially true for people who hold positions of supervisory responsibility. When dealing with innovative storage and networking paradigms, such as network virtualization and cloud computing, monitoring becomes a far more difficult function to do than it would be under normal circumstances. This is due to the rising mutability of the underlying networks and resources, in addition to the greater diversity of those networks and resources. The reason for this is because the underlying networks and resources are becoming more diverse. This is because new storage and networking paradigms have been created, such as network virtualization and cloud computing, as was indicated in the paragraph that came before this one. These new paradigms have made it possible to accomplish the goals that were before unachievable. As a direct consequence of this, this occurrence has taken place. This is only one example of a reason that has contributed to the precipitous growth in popularity of cloud computing over the course of the last few years, and it is just one of the many variables that have played a role in the meteoric rise in popularity of cloud computing. Because there are currently no industry standards in place for monitoring cloud-based information infrastructure, the process of reaching a consensus on how to evaluate cloud-based information systems is made more difficult. This makes it more difficult to come to a decision on how to evaluate cloud-based information systems. As a consequence of this, achieving a consensus on how to evaluate these sorts of systems is a process that is plagued with a great deal of challenges since there are so many various methods to carry it out. As a result of this, the process is laden with a great deal of difficulty. An approach that is layered and operates from the top down is the most effective method for understanding the complexities of monitoring cloud services, which range from the underlying physical and virtual infrastructure all the way up to the levels of the end-user and service provider. The best way to understand these complexities is to use an

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

approach that employs this method. The technique that is being described here is referred to as the top-down, layered approach. This is the most efficient method for completing the tasks that need to be completed. The most beneficial course of action that one may take in this situation is to react to the situation in the way described above. In order to conduct assessments in the areas of dependability, security, compliance with SLAs, and efficiency, the metrics and criteria that are incorporated into each layer of measurement are used. Because of the highly competitive nature of the cloud services business, the leading market representatives in the sector have been divided into a broad range of categories determined by the kind of services that they provide to their respective clientele. After we have completed this stage of the process, we will move on to the subsequent stage, which entails conducting an analysis of the management strategies that are specific to each service category, and then transferring those strategies into the appropriate operational levels. This stage will be reached once the previous stage has been completed [37].

According to the National Institute of Standards and Technology, cloud computing permits the rapid provisioning of scalable, elastic, and low-overhead networked resources on demand. This is made possible by the cloud's decentralized nature. This is something that can be accomplished in a very little period of time (NIST). There is a possibility that using cloud computing will bring a number of major advantages, some of the most important of which are increased productivity, scalability, speed, and accessibility. The benefits of using cloud computing are extensive, with the most important advantages having been covered up to this point in the discussion. The use of cloud computing is not only highly beneficial in terms of saving money, but it is also very useful and simple to put into action. The service for cloud computing, similar to other types of services, does not come in a format that is appropriate for each and every user. Even if there are several cloud computing service types now accessible, this is still the case. As a consequence of this, it is essential to provide a vast array of service models and solutions in order to suit the needs of a large customer base that you are servicing. If you do so, you will ensure that your business will be successful. Putting an end to this problem may be accomplished in a myriad of various ways, each of which is a legitimate alternative in and of itself. There are already a large number of different options available that may be used in order to get access to the cloud and make use of the resources contained within it. IaaS, PaaS, and SaaS are the three basic categories of cloud computing service models. These three primary categories are also referred to by their respective acronyms. "Software as a Service" is referred to using the acronym "SaaS." (AaaS). Models such as "Infrastructure as a Service," "Platform as a Service," and "Application as a Service," abbreviated "aaS," are just a few examples of the many different kinds of things that may be employed in a circumstance such as this one. There are many more possibilities as well (SaaS). They are inextricably linked to one another, and the whole thing is what's known as the "cloud service stack." It is difficult to engage in any activity if even a single one of them is missing. The term "cloud computing" is becoming more common as an increasing number of individuals and businesses become aware of the numerous benefits that can be obtained by making use of the services that are made available by cloud computing. This is one of the reasons why the phrase "cloud computing" is becoming more well-known. There will be a growth in the overall amount of

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

information that is accessible as a result of the increased number of people who are utilizing them owing to the fact that more and more people are using cloud services. The completion of the project will be an undertaking that has to be conducted with a higher degree of difficulty as a direct result of the turn of events that has taken place. In light of this fact, it is of the utmost importance to have solutions that are at the cutting edge of what is currently feasible in terms of cloud administration and monitoring [38].

The relevance of making sure that one is alert while monitoring both the services and the infrastructure at all times. If you want to keep a watch on cloud computing, you are going to need a hybrid strategy that incorporates both automated monitoring tools and human monitoring tools. This will allow you to keep an eye on cloud computing. This is a very important consideration. The measuring, monitoring, and administration of activities that are now occurring inside an architecture that is based on cloud computing are the primary goals of cloud management. Cloud management is based on cloud computing. The fact that cloud management takes place in an online setting is one of the characteristics that set it apart from other methods. Using either the platform or the technology that is made available by the cloud, testing may be carried out manually or automatically, depending on the desire of the individual doing the testing. Both strategies ought to be taken into consideration. Calculating difficult mathematical equations in the scientific field or managing cloud resources that are critical to the operation of a firm both need the presence of this vital component, which must be present in order for either of those things to be possible. A significant portion of the obligations that are associated with the cloud may be made a great deal easier with the help of cloud administration [39].

Statistics and financial transactions need to be coded. The CMP demands meticulous record keeping in order to prevent falsification and manipulation, while cloud computing necessitates the use of both virtual and physical resources in addition to resource allocation. This is done to ensure that sensitive information is maintained in strict confidence at all times. Both of these preventative measures are carried out in order to ensure that the data that is being kept retains its original form [40].

Included in the terms of service contract are clauses pertaining to QoS, pricing, and penalties. The information that pertains to CMs, resources, and SLAs will be gathered by the Parameter Monitor so that it may be analyzed for the purpose of carrying out service verification once this information has been gathered. Tailoring cloud services to satisfy the specific requirements of each and every unique customer The CMT is the one who is in charge of keeping a careful check on how all of the available resources are being used as well as documenting the data. The assistance offered by this allocation and reallocation tool has the potential to result in the implementation of allocation strategies that are more advantageous in general, which would be a positive outcome [41].

It gives the certainty that cloud-based firm operations can continue to function at their usual levels of efficiency, which is a vital component of capacity planning. [Capacities planning] The CMT has the ability to identify the areas in which service providers are either failing to properly manage their resources or spending an excessive amount of money on those resources. It is a

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

collection of guidelines that explains how applications and computer programs need to carry out the responsibilities for which they were developed. This function ensures that any modifications that users make to their use of the cloud, such as adding or deleting resources, are rapidly reflected in the services that those users make use of. As part of their responsibilities, CMTs are accountable for configuring the services in their entirety, which includes all of their individual components. When dealing with sensitive information or using services that are located in the cloud, it is extremely important to give 100% assurance that your data is secured from unwanted access. This is one of the most important aspects of cloud computing [42].

The exponential development in the amount of data that is being held in various places around the globe is posing problems for the information technology sector. These problems are a direct outcome of the exponential growth in the quantity of data. The entire quantity of data in the universe is expected to expand to 175 zettabytes by the year 2025, up from 33 zettabytes in 2018, according to an estimate that was made in the white book named "Data-Age 2025," which was authored and published by International Data Corporation. It is anticipated that the field of cloud computing would see great expansion over the course of the subsequent number of years. Cloud storage is one strategy that may be taken to solve the difficulties that have surfaced as a consequence of the exponential expansion of data. Cloud storage is a term that refers to the act of storing data via the internet. The concept of cloud storage refers to the process of combining several storage devices into a single storage pool by using a network that is linked to various data centers. Users are able to upload data to a shared server utilizing cloud storage, from which they may afterwards retrieve files using any device that is connected to the Internet at any time and from any place. Users are able to do this regardless of where they are located. Users that make use of cloud storage have the ability to do this retrieve [43].

The term "cloud computing" refers to a way of storing data and retrieving that data that depends on distant servers rather than on-premises technology. The term "cloud computing" is also used to characterize the approach. The technique itself is often referred to by the name "cloud computing," which is another usage of the word.The choice that was taken by the founders of the company to move its operations to the cloud will have a significant and positive impact on the firm's clients in a variety of different ways, each of which will be helpful in its own manner. A few of the many benefits that may be had include scalability, dependability, high performance, availability, and pricing that is based on a pay-as-you-go model. These are just some of the many advantages that can be enjoyed. When services are subcontracted to other parties, there is an increased risk of sensitive information being stolen, data being lost, and necessary services being denied [44].

We have absorbed a variety of well-known cloud storage services, such as OpenStack Swift, Ceph, Dropbox, Google Drive, and Microsoft OneDrive, amongst others. There are several more options available for cloud storage, including Amazon S3, Google Drive, and Microsoft OneDrive. This emphasizes the fact that the cloud computing industry is now facing its most significant issue, which comes in the form of concerns over the security of data. Customers want their cloud service providers (CSPs) to protect the data that they make available to them,

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

ensure that the services they make use of will continue to be dependable, and demonstrate that they comply with all applicable security regulations [45].

Because the aspect of security that is being evaluated is intangible, it may be difficult to determine the level of protection that is offered by a CSP. The process of consumers embracing cloud computing is sped up since they are better equipped to make informed judgments about the security of the services they are contemplating hiring. Customers are able to adopt cloud computing in a more expedient manner as a result of this. The end effect is a higher rate of adoption in a shorter amount of time than was first predicted, and this is the case since the timing was better. The limits of the connection between security service providers and their customers in terms of security management are now being outlined in the form of a Security Service Level Agreement, which is currently being drafted by both parties. This is being done to guarantee that customers are provided with an adequate level of protection. This action is being done in order to protect the interests of both parties, and we apologize for any inconvenience this may cause (also known as a Security SLA). The interests of all parties are safeguarded by this agreement, and all parties are required to cooperate with one another in order to forestall breaches of security and the accompanying monetary and legal repercussions. Within the confines of this contract, the interests of all parties are safeguarded. As a direct and immediate consequence of the idea of cloud computing, cloud service providers, also known as CSPs, are faced with a broad variety of challenges. These challenges may be broken down into many categories. Guaranteeing production, restricting resources, planning for calamities, allocating tasks across regions, and adhering to legal requirements are some of the most important challenges that organizations need to overcome. In an effort to find a solution to the issues that have been cropping up, the idea of federating many cloud services was devised as a possible course of action. It makes it possible for a CSP to pass over some of the requests made by its users to third-party service providers while at the same time keeping complete control and visibility over the way in which those requests are carried out [46].

The following is a list of categories that might be used in the process of classifying cloud services:

Software as a Service, which is more often referred to as SaaS and stands for Software as a Service, is a distribution method that makes advantage of cloud computing to provide a diverse group of users with access to a collection of helpful software tools. The term Software as a Service may also be shortened as SaaS. "Software as a service" is shortened to "SaaS," and the full acronym is "SaaS." Users are required to employ a web browser in order to first check in to their individual accounts before being allowed to make use of the applications. They will charge you solely for the amount of time that you spend using their software, which includes Microsoft Word, Notepad, and Paint, among other programs. Software as a service (SaaS) is offered by a number of the most well-known companies in the cloud computing business, including Google, Zoho, Intuit, and Salesforce.com, to mention just a few. These companies are also leaders in the cloud computing market [47].

As a direct result of the common practice of providing end users with a Platform-like Operating System (OPS) in the form of a service, the term "Platform as a Service" (also written as

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

"Platform as a Service") came into being (PaaS). Users of Integrated Development Environments (IDEs) have the opportunity to construct their own frameworks with the assistance of a number of different tools, including a compiler, editor, and debugger, amongst others. The IDEs are responsible for making this a reality. PaaS suppliers include organizations such as Google Apps, Bungee Connect, and Force.com, to name just a few examples of the kind of companies that fall under this category.

IaaS, which stands for "Infrastructure as a Service," is a paradigm for the delivery of cloud computing services that involves outsourcing the administration of various hardware components to a third party. This concept was developed by Amazon Web Services (AWS) and Microsoft Azure. Servers, networks, and storage are all examples of these types of physical components. Customers use the services in a way that is compatible with the specific demands that are put upon them by the budgetary limitations that they are under. The user is in charge of the administration of the operating system as well as the deployment of the application; nevertheless, the user does not have any authority or control over the cloud infrastructure at their disposal. One company that offers its customers infrastructure as a service in the form of instanced hosting is Amazon, which is just one example of a company that does this. Elastic Compute. Concerns over the localization of data as well as the integrity of the data are two potential impediments to the broad adoption of software-as-a-service (SaaS) applications. The majority of the time, the customer is not informed of the location at which the service provider retains its data or the precautions that must be made to prevent it from being altered without their permission. Additionally, the customer is typically unaware of the security measures that must be taken to protect the data from being compromised. The loss of the level of trust that already exists between end customers and the cloud service providers that those customers work with as a result of the use of SaaS is one of the most serious problems that could develop as a result of the use of SaaS. This could be one of the most significant issues that arise as a result of the use of SaaS [48].

The IT team at the institution may decide to host the SaaS application on a dedicated server or to make use of the infrastructure services provided by reputable third-party suppliers such as Google, Amazon, and others in order to guarantee that the personal information of students is kept private. This choice may be made in order to protect the students' right to privacy. Both of these choices provide the educational organization the capacity to safeguard the personally identifiable information of its pupils. Which of these two policies, if any, will be implemented at the school is a decision that must be made by the administration. Because of all of these various factors, the vast majority of secondary schools that took part in our survey have decided to use private clouds rather than public or hybrid varieties of cloud computing. This decision was made because private clouds offer more security and privacy than public or hybrid clouds [49].

When using a platform known as PaaS, software engineers who work for educational institutions may have a better chance of constructing and moving SaaS apps to the cloud in a more expedient way. This might be the case if they are able to do so more quickly. The term "platform as a service" is abbreviated as "PaaS." PaaS has the ability to make both of these

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

duties easier to handle, which is why this choice is available to users. However, there is a possibility that developers would have difficulties whenever they use PaaS services. This is something that need to be anticipated and prepared for. An increase in capacity makes it easier for programmers to coordinate the operations of third-party cloud services, which is essential for getting things off the ground. [This has to be explained in more detail] Using a PaaS comes with a number of significant drawbacks, one of the most significant of which is the possibility that customers may be restricted to using only specific programming paradigms or tiers of service providers. This is just one of a number of significant drawbacks that come with using a PaaS. Using a PaaS comes with a lot of important downsides, including this one, which is only one of them. Using a PaaS has a number of extra significant drawbacks that come along with it, and these drawbacks are significant in and of themselves. This is only one of several significant issues that are occurring all around the globe at the present time. Everyone involved will need to give these models and services a complete overhaul and rethink them from the bottom up before they can be transferred to an environment that is compatible with PaaS. This must be done before the models and services can be moved to an environment. They need this since they are very dependent on the circumstances that they find themselves in at the present. It is difficult for users to move to one of the other available alternatives since they are only allowed access to the network to which they are now connected. This makes it possible for consumers to transfer networks. Developers are able to create and maintain applications on top of the platform, but they are uninformed of the degree of security that the platform's service provider has permitted to be present in the underlying infrastructure. Despite this, developers are still able to create apps that run on top of the platform and maintain such applications [50].

When compared to the other two types of services, customers whose data is protected by an IaaS provider have a greater degree of control over the level of security that is applied to their information. This is because the IaaS provider is in charge of maintaining the infrastructure that underpins the security measures. The dependability of the information that is offered by the source is by far the most significant factor to take into consideration when picking a source. In a setting that makes use of infrastructure as a service, also known as IaaS, the responsibility for ensuring that the system continues to function properly rests in equal measure with both the customers and the service providers. Moreover, the customers bear the primary responsibility for ensuring that the system continues to function properly. Because giving their services is a duty that comes with it, service providers are obligated to take severe safeguards to protect both the physical and digital environments in which they operate. This is the case because providing their services is a responsibility that comes with it. These prophylactic steps have to be taken at the same time as one another. When using cloud storage, it is the user's obligation to ensure that their own data, apps, and computer operating system are protected from harm (OS). The concept that is referred to as "Infrastructure as a Service," or IaaS for short, requires virtualization in order to function properly. Virtualization is important. Without it, the idea would not be able to function as intended in any meaningful way. When a number of users are using the same hardware in a virtualized environment, there is always the risk that an attack might be carried out concurrently against a number of distinct tenants. This is because several users are utilizing the same hardware. In order to successfully compromise the overwhelming

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

majority of cloud accounts that are associated with tenants, an adversary would need to have access to the root level of the cloud [47].

The use of cloud computing has made it easier to carry out some obligations; nevertheless, it has also introduced new security threats that need to be taken into account. Cloud computing may make certain responsibilities easier to carry out. When data is held in a variety of cloud services and is dispersed throughout a network, there is a significant possibility that there will be a great number of feasible entrance points that bad actors may exploit. This is because there are more entry points, the more likely they are to be exploited. This is because there is a higher number of sites where data may be held, which is the primary reason for this. A user-oriented security model, data security, network security, application security, and security rules are the five fundamental components that need to be present in a cloud environment in order for such environment to be regarded as secure [51].

Security policies provide an in-depth explanation of the processes that need to be carried out and the security rules that need to be adhered to in order to prevent a system from being hacked. These must both be done in order to prevent a system from being compromised. These recommendations were formulated with the intention of ensuring that the processes in issue are adhered to at all times, irrespective of the conditions that may be present. If you adhere to these guidelines, you should be able to build a cloud workspace for your company that is reliable and secure. Concerns over client management, service level agreements (SLAs), previous levels of trust, and other factors may have an effect on the security measures that are finally put into place, which raises the possibility that these factors may influence the security measures in some way [52].

An agreement that is known as the Service Level Agreement governs each and every one of these transactions that take place between the customer and the business that is in charge of providing the service. This agreement is in place to ensure that the customer is satisfied with the level of service that is being provided (SLA). Service level agreements are a helpful tool for providers that are worried about their reputation and who make every attempt to live up to the expectations that are established by their customers. This kind of supplier makes every effort to meet or exceed their customers' expectations (SLAs). On the other hand, in line with the provisions of service level agreements, service providers are often released from their responsibilities in the event that mistakes are made or outcomes are not up to standard (SLAs). When doing an examination of the overall quality of a service, it is very important to take into account the service level agreements (SLAs), which are more often referred to by their acronym. This is because the SLAs outline what is considered acceptable performance for the service. On the other hand, because to a lack of adequate knowledge, the SLA is unable to provide assurance at this time for a specific solution that has been proposed. To set the record straight and dispel any misconceptions, the Service Level Agreement (SLA) does not guarantee excellent service and is unable to change the quality of the delivery if it does not meet the parameters that were expected. The SLA includes not only a statement of its worries about the resources that are subject to its inspection but also a list of the resources that are subject to its examination. The statement of its concerns about the resources that are subject to its inspection

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

may be found in the SLA. In addition to this, the SLA provides a statement that outlines the concerns that it has about the resources that are within its purview of inspection. In addition to this, it outlines the responsibilities that are held by both the customer and the provider. These duties include both the obligations that are owed by the supplier and the expectations that are owed by the client. According to the terms of the Service Level Agreement (SLA), the provision of the services must take place within the allotted span of time (SLA) [53].

The ability to do computing tasks in the cloud gives businesses the chance to save costs while at the same time expanding their data storage capabilities. As a consequence of this, the disclosure that service providers are now permitted to make use of novel approaches to the verification of customers serves as a source of trust as a result of the increased level of assurance it provides. This is because of the fact that customers are given the opportunity to use these novel approaches. Depending on the conditions of the procedure, the amount of authentication carried out by computers may vary from full to partial verification of the user's identity. In the context of cloud computing, the use of a simple authentication method will not, on its own, get you access to the information that is stored with the different cloud service providers. It is still feasible to accomplish this goal, despite the fact that doing so carries the potential danger of disclosing sensitive information to individuals who have been granted increased rights. The undeniable fact that doing so presents a challenge does not alter the fact that this is, in fact, the case. In spite of this, there is a broad variety of additional authentication methods that can be used to get around this issue. These methods can be used to circumvent the difficulty. These methods may be used to avoid having to deal with the problem. You are able to circumvent this challenge by using the solutions that were discussed before. If you follow these steps, you will be able to prevent yourself from coming across this problem. When a user offers information that is only known to them, such as a card number or a password that they have selected themselves, the system will often provide that individual with exclusive access to the resource that is being inquired about. A card number or a password that the user has concocted on their own would be an illustration of this kind of information. Take for instance a card number or a password that they have chosen on their own own as an example. These processes may be categorized, in the overwhelming majority of cases, as belonging to one of the two primary categories that are outlined in the following paragraphs: The two most fundamental forms of security verification are known as authentication via the use of the user's physical assets and authentication through the use of digital means. The usage of the user's actual possessions in the authentication process is becoming increasingly widespread. The authorization aspect of information security refers to the section of the system that is responsible for determining whether or not to grant a user's request to get access to a certain resource. If the request is granted, the user will be able to utilize the resource in question. In the event that the user's request is accepted, access to the resource in issue will be allowed to the user. This option is ultimately a consequence of the duty that they have taken on. In the case that the request is honored, the user will be provided with access to the resource that was the subject of the inquiry. When acting in this position, the administrator acts as a communication channel between the system and the individuals who use the system on a day-to-day basis. This allows for a more efficient flow of information. Because the cloud is not a

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

centralized location, a single client may utilize the services of many different providers, each of which may take a slightly different approach to the safety of the client's data. This is because the cloud is decentralized. Because of how the cloud works, a single customer may make use of the resources provided by a diverse collection of service providers. This is because the cloud operates in a distributed manner, which is the explanation for this outcome. This is the reason for this result. After the user has given the software permission, one of the potential effects is that the program will then be made accessible to the general public. This is one of the results. This is one of the outcomes that may take place. Access management rules or access delegate credentials are the two options available to users in this specific setting. They have complete autonomy over their choice. They have unrestricted discretion over the option they choose. These functions are only accessible to and used by the persons who have been granted authority to make use of the system's administrative controls and operational capabilities. Other users are unable to access or utilize these features [54].

A centralized access system not only improves the security of the data to which it allows access, but it also makes the administrative and security procedures that much more efficient. This is because the data to which it grants access is protected from unauthorized access. Another benefit is that you will not waste as much time as you typically would by doing things that are not necessary. There is more than one kind of authorisation system, the most prevalent of which are the MAC, DAC, RBAC, and ABAC. Other types of authorization systems are also possible. The MAC is the one that sees the most action among them all. In addition, there is a diverse selection of authorization processes available for your selection [55].

Every single one of the files that make up the logical pool is copied to a digital archive that is kept in the cloud, and this archive contains all of the files. The files are preserved in this archive, which may be accessed here. The information is scattered among a very large number of servers, each of which is managed by a separate company that provides web hosting services. There is the possibility that each server might be situated in the real world under its own unique localization. Distributed computing systems will need data storage that is not only safe but also simple to access as the number of people who use the internet and the number of items that are linked to it continues to grow. This is due to the fact that both the number of people who use the internet and the number of items that are linked to it will continue to increase in the foreseeable future. This is due to the fact that both the number of people who use the internet as well as the number of gadgets that are linked to it are continuously rising at a fast pace. As a direct consequence of the challenges that are linked with data management, concerns have developed over the storage, availability, and confidentiality of data, as well as the CIA and many other types of security. In addition, there are many more forms of security. If you want to store your data in the cloud, the organization that is in charge of administering that cloud should also be the entity that is responsible for guaranteeing that the data can be accessible at any time [56].

When it comes to cloud computing, there are three key challenges that need to be addressed and resolved: availability, transparency, and safety. Availability refers to the ease with which a service may be accessed by users. (CIA) Any and all data that is stored in the cloud should

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

ensure that it conforms to the ACID property in order to secure data privacy and preserve the ability to track data. Regardless of the kind of data being used, this step has to be taken. Not only is it advantageous for providers of cloud services to have a high degree of availability for themselves, but it is also advantageous for the consumers that they serve to have such availability. Hardware or software that isn't performing as it should, hostile assaults from the outside, or both of these things happening at the same time are just a few of the many different reasons that could put the availability of a service in jeopardy. Other possible causes include: OneDrive, a cloud storage service that is supplied by Microsoft, does not provide any type of encryption for the sensitive information that its customers upload on the site. Dropbox and Google Drive, on the other hand, both provide cloud storage services and provide encryption for sensitive data that their customers save on their own websites [57].

Simply said, metadata is a collection of files in their most fundamental form. This is the definition of metadata at its most fundamental level. The development in popularity of cloud computing as a preferred form of computing has corresponded with a simultaneous rise in both the relevance of metadata and the degree of complexity that it includes. This rise has occurred at the same time that cloud computing has become more popular. The database that was once used for the storage of sensitive information is now undergoing a transition that will see it function instead as a repository for metadata. The term "metadata" refers to information that may be used to describe a wide range of things, including who did something, where it was done, how it was written, and how the data was stored. It's conceivable that this information will be kept in a database for future reference. It is possible to save the information in a number of various file formats, all of which are selectable by the user depending on their preferences. It is important to keep in mind that there is a potential that some of the information may provide more explanation about the structure of the data. This is something that should be taken into consideration. Cybercriminals have the capacity to collect metadata, which may include information of considerable value and which they may obtain access to and use for their own benefit. Cybercriminals have the ability to get their hands on metadata. On the other hand, companies are likely to start using metadata in order to get an even higher monetary value from the data that they are currently in possession of. This may be accomplished by using the data to make predictions about future events. In addition, it is vital to encrypt metadata in a way that is acceptable since metadata may include information that has to be kept secret or is sensitive in nature. Unfortunately, only the massage records may be encrypted; as a direct consequence of this constraint, the touch data can still be accessed by anybody who has an interest in doing so. When one makes use of a virtual private network, which is also often referred to as a VPN, encrypting communications that may include potentially sensitive information is a breeze. This kind of network is also sometimes referred to as a VPN. A virtual private network, or VPN for short, is one of the many names that may be given to this kind of network [58].

If you are concerned about the safety of your data while it is being stored in the cloud, one of the most significant threats that you should be aware of is the presence of vulnerabilities within the application security. If you are concerned about the safety of your data, you should be aware of this risk. When developing software applications, there are a variety of distinct elements of application security that need to be taken into account. These issues include: This is particularly

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

true with regard to the installation of software and networking that is hosted on the cloud for usage with applications. Providing web-based business services that can be relied on is of the highest significance. The primary purpose of the Open Web Application Security Project (OWASP) is to improve the amount of security that is already present in the infrastructure and server-side code that are used in online applications [59].

Despite the fact that this is the case, there are still a great many obstacles to overcome in order to ensure the safety of web apps against vulnerabilities. Since mobile applications have been one of the most frequent ways for people to engage with the web for quite some time now, we can safely say that they are one of the most popular methods overall. In point of fact, hackers have the ability to insert hazardous executable malware into communications that are sent by endpoints, which may result in damage being caused. It's feasible that this will put our safety in jeopardy in some way. When talking about computer security in a technical manner, this kind of attack is referred to as "code injection," and that is the word that is used to describe it. Another method that may be used to inject harmful code into a website is called cross-site scripting, and it can be utilized by hackers. Using this method, one might potentially get confidential information. This method is also known as cross-site scripting in certain circles (XSS). Because of this particular characteristic of their composition, the scripts may be differentiated from one another. Inadequate or faulty coding in the threshold, exception, or password protection systems of online applications is one potential cause of vulnerabilities that might exist in these programs. When there are problems with management, they are often the consequence of improper setups or alterations that were not carried out in the right way. This may also be the case when there are problems with other aspects of the organization. Because of this need, programming languages that are used for certain reasons often have extra barriers included into the design of the language itself. Taking precautions to secure online applications will make it more difficult to handle other interconnected difficulties, such as guaranteeing that the network's security will continue to be maintained with absolute integrity. Operating systems are depended upon to a great degree by a variety of software, including cloud-based security software, application software, and the operating systems themselves. It's probable that this reliance should be considered carefully. It is the responsibility of the operating system to ensure that everything continues to function correctly despite the increased demand that occurs when multiple programs want access to the same piece of hardware. This increased demand can be attributed to the fact that multiple programs are competing for the same resource (disk space, RAM, display real estate, etc.). This takes occur when many applications all request access to the same piece of hardware at the same time [60].

There have been concerns expressed over the safety of user information that is kept in the cloud, which is due to the fact that cloud computing is, at its core, a network-based model. The similarities between network security and cloud security may, in a fundamental sense, be separated down into three areas: generating data protection, network security, and the information security sub-area. Developing data security is the most basic of these commonalities. In other words, you may see these three categories coming together to make a triangle if you put your mind to it. The potential threats to network security that might come from the outside world are varied and extensive in the spheres in which they could have an

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

impact. Administrators of networks are obligated to comply with all applicable security requirements whenever they use data and cloud technology safeguards and services. These restrictions may be broken in a number of ways. This is the case everytime they deploy these technologies [61].

It is discouraging to learn that there have been fresh problems with connection appearing recently. Some of the new issues affecting connection include denial-of-service attacks, distributed denial-of-service assaults, flooding attacks, and vulnerabilities in network protocol. In order to defend oneself from possible threats, the technique that is both the easiest to implement and the one that is the most successful is the installation of firewalls. The involvement of end users, servers, and cyberspace routers as potential audit targets may be required for audits of the data that is kept in the cloud. In terms of their operating capabilities, there is no noticeable difference between traditional firewalls and firewalls that are supplied in the form of a service. Conventional firewalls have been around for a long time. They have the ability to access the service via any network of their choosing since it is hosted in the cloud. This gives users a great deal of freedom [62].

A firewall as a service, also known as FWaaS, is an intriguing option for businesses of any size who are concerned about the upkeep of their network security since it exceeds on-premises firewalls in terms of cost, performance, and flexibility. Despite the fact that they are not perfect, firewalls have the potential to be useful in improving communication and reducing the risk of an attack, both of which are important factors in ensuring the safety of cloud services. Despite the fact that there is no such thing as an impenetrable firewall, this is nonetheless the case. It is feasible for a firewall to prevent malicious software from entering the system, such as backdoor Trojans; but, a firewall is unable to remove viruses, worms, or any other forms of malware that are already installed in the system. Communication is essential for a firewall; a firewall only examines packet headers when it detects and logs malicious network access; a firewall must take into account protocol sorting, protected addresses, destination addresses, and a secure port policy; and a firewall must only examine packet headers when it is necessary to do so. Communication is essential for a firewall. Communication is necessary for a firewall to function properly. In order for a firewall to perform its intended functions, communication is required. IDS can not only stop undesired intrusions, but it can also gather data that may be used in the future to make a system more secure against potential threats. This data can be used in the future. The phrase "unnecessary network access" is shortened to IPS, which is an acronym. The abbreviation is used to denote the meaning of this phrase. When it makes that statement, it is referring to just this point in particular. The purpose of intrusion prevention systems, sometimes abbreviated as IPS, is to recognize and monitor any possible threats that may arise. On the other side, they may also be effective in awarding against invasions and ransomware. This is another potential use. In conclusion, it is very important to bear in mind the larger meanings that are associated with the phrase "network security." Extra reading material [63].

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

2.1 GPU Architecture History and Evolution

The production of the modern graphics card began in 1995 with the introduction of the first 3D add-in cards. From the perspective of parallel architecture, there are three generations

A. Fixed functional architecture: Each hardware unit has graphics between 1995 and 2000 pipeline processing functions are fixed in this generation, many. On processor cores executes all operations same on the stream. Representation. Utilizing the graphics processing unit or GPU, a great deal of graphics processing can be facilitated.

B. separated shader architecture: The GeForce 3 from NVIDIA brought programmable pixel shading to the consumer market in 2001. This helped greatly increase the expressiveness and versatility of fast graphics processing. Usually used in games.

C. Unified shader architecture: In 2012, NVIDIA launch

Kepler technology that use dynamic parallelism. The Pascal architecture was implemented for higher performance computing. In 2016, with the advancement of AI, Deep Learning, autonomous driving systems, and numerous other computer-intensive applications, NVIDIA released Pascal to enhance performance.

2.2 GPU Computing Architecture (Modern).

This illustrates not just the functional power of a modern graphics processing unit, but also its visual appeal. A graphics processing unit, often known as a GPU, is constructed from a large number of individual processors that are referred to together as streaming processors. These streaming processors each have their own GPU cores, which are also sometimes referred to as streaming processors, in addition to specific function units, registers, double precision unit(s), and a thread scheduler of their own. Streaming processors are also sometimes referred to as graphics processing units (GPUs). A graphics processing unit (GPU) is often responsible for graphics processing. In addition, the graphics processing unit, also known as the GPU, is accountable for carrying out the bulk of the program instructions via the utilization of floating point arithmetic. The graphics processing unit (GPU) is in charge of multithreading, and this component may manage anywhere from 32 to 96 threads at the same time, depending on the technology. Device memory and chip memory are the two levels of memory that are present in graphics processing units. These two layers make up the RAM that is available on the GPU as a whole. When reading and writing data to and from the device's internal memory, a method known as dynamic random access is used. In addition to the local, global, and constant memories that are provided by DRAM, on-chip memory includes a shared memory register, an L1/L2 cache, a constant cache, and a texture. This is in addition to the memory that is provided by DRAM. Additionally, the memory that is contained inside the chip has a texture. The working memory is managed by the central processing unit, and it is divided up and distributed throughout the several threads that make up the program. The logical constant and the texture memory were both constructed from the ground up with the intention of carrying out graphical calculations as their primary task once they were brought into existence. This intention guided their creation from the very beginning. It is not unusual for hundreds of clock cycles to pass

ISSN 1533-9211

DOI 10.17605/OSF.IO/7E5Q9

before the memory that is physically present on the device itself can be accessible. This delay is due to the fact that it takes time for the memory to be read from and written to. In a perfect world, we would want to shorten the length of time, on average, that clients are required to wait before receiving service by as much as we possibly can [64].

There are a great number of unique iterations of cache levels L1 and L2 available. When caching is not necessary, such as in the case of the processing of a stream of data, graphics processing units, which are sometimes referred to as GPUs and are regularly used in circumstances in which they may be employed, are frequently used in situations in which they may be utilized. Because each SM has its own speedy but limited shared memory (48 KB in the NVIDIA K40 GPU3), the memory can only be accessed by threads that were first established on that specific SM. This implies that the memory cannot be shared across SMs. Only these threads have permission to access this memory location. It cannot be accessed by any other threads. (This is the primary reason why the graphics processing unit (GPU) on the chip only supplies a very modest amount of space to the cache.)[5]. In addition to the main central processing unit, the graphics processing unit, which is more often referred to as the GPU, is frequently utilized in the role of a secondary processor. The communication protocol that is utilized between a graphics processing unit (GPU) and its host is known as PCI Express. Data may be automatically transferred between locations on a computer if it is equipped with both a graphics processing unit (GPU) and a central processing unit (CPU) on the same circuit board (DMA). The ability to deliver a zero-copy function is only available to modern GPU architectures like CUDA and OpenCL. Because of this function, mapping addresses from the host to the system is now both possible and practical. In order to successfully complete this mission, this strategy places a strong focus on physical closeness, which makes communication much more effective[65].

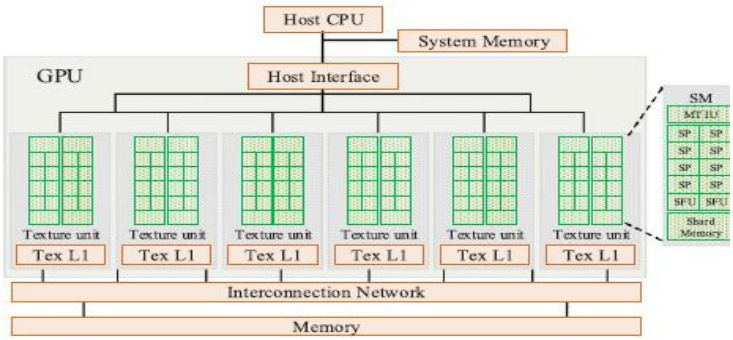

Fig. 1. GPU Architecture [66]

2.3 GPU APIs and Programming Models

Because OpenGL, Direct3D, CUDA, and OpenCL have been the primary targets of API remote and other GPU virtualization techniques, these application programming interfaces and programming models are being presented as GPU application programming interfaces. This is due to the fact that these four APIs are being presented as GPU application programming interfaces (APIs).

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

In the year 1999, Woo and several of his colleagues developed what is now known as the OpenGL graphics acceleration library. It confers onto software applications the capacity to make advantage of the graphics processing unit (GPU) of a computer (GPU). The library was developed in such a way that it would facilitate the construction of computer games, the processing of images, and the presentation of programs that are used in technical fields in a manner that would be easier to do. OpenGL is an application programming interface (API) that has received significant popularity because to the fact that it is independent of the hardware on which it is used and does not need special support[67].

Regardless of the platform that they are designed to function on for the hardware, video cards often provide graphical capabilities as a component of the feature sets that they offer.

According to Blythe (2006), the graphical application programming interface that is known as Direct3D is used by Microsoft Windows. [Further citation is required] The genuine article is included inside this Direct3D content package. Authenticity is always guaranteed under all circumstances. It is a low-level application programming interface that is used to produce 3D graphics in applications that need a lot of resources, such as games. One example of such a software is Blender, which is used to create virtual reality experiences. The creation of animations is one application that may make use of this interface. On top of the complexity that the GPU hardware implementation already has, Direct3D provides a layer that is both standard and universal. This enables the usage of more complex graphics capabilities like Z-buffering, W-buffering, stencil buffering, and spatial anti-aliasing, among others. Microsoft was the company that first created Direct3D. Both Microsoft Windows and Microsoft Windows Vista include support for the Direct3D graphics library [68, 69].

CUDA is a programming paradigm that was developed by NVIDIA [NVIDIA 2007b]. It was built with the intention of being used in GPU-accelerated parallel computing. To put it another way, it makes it possible for programmers to use graphics processing units (GPUs) that are outfitted with the CUDA parallel computing platform in order to process data in parallel for a wide variety of applications. These GPUs can do this because of the CUDA parallel computing platform. These graphics processing units (GPUs) may be put to use as long as the CUDA parallel computing framework is loaded onto them [70].

GPUs (graphics processing units) that are designed to be used by a large number of people (GPGPU). CUDA is a programming environment for parallel computing that has strong links to the computer languages C and C++. CUDA was developed by Nvidia. Both of these languages have had a limited set of primitives added to them so that they may be used to manage system memory, send data, operate the GPU kernel, handle events, and carry out atomic and synchronized operations. These fundamental building blocks are compatible with the programming languages C and C++ [71].

OpenCL is a technology that can be used to build parallel programs that are capable of operating on a variety of different computer systems. You may utilize this technology to design parallel apps using OpenCL. In addition to that, it defines a programming language for the purpose of constructing compute kernels. This language is called OpenCL C, and it is pretty

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

similar to the language C. The objective of this language is to create compute kernels. [Group et al., 2008] The application programming interfaces (APIs) that are made available encompass not only the procedure of loading kernels into an OpenCL device but also the control of memory transfers between the host and the device. It is not possible to utilize CUDA with OpenCL since NVIDIA is the only company that makes CUDA available to the public. This is due to the fact that NVIDIA is the only company that makes CUDA accessible to users [72, 73].

2.4 Graph Programming Models

The vast majority of contemporary GPU-based graph processing frameworks make use of a programming interface that is vertex-centric. This instantly iterates the feature over all vertices in order to boost the amount of programming freedom that is accessible to its utmost possible level. In order to make modifications to graphs in parallel, there are now two new models that may be employed [74].

GAS and BSP are the two key notions that need to be taken into account, so keep them in mind (Bulk Synchronous Parallel).

The acronym A. GAS Hypothesis is often used to refer to this theory. The contention that GAS is an efficient simulation model for graph algorithms is one that can be supported with evidence. [Further citation is required] The functions of collecting, applying, and dispersing are the ones that are most often carried out in an implementation that is vertex-centric. As a result of this, the GAS model may be broken down into the following three phases [75]:

The process of gathering together a number of separate people into a single group. When we get to this point, a vertex will make use of the collecting function that the user has given to it in order to collect data from other vertices and edges that are present in its immediate vicinity. This will take place when we reach this stage. — making a request to take part in something by submitting an application. At this step, we make use of the information that we acquired in the previous phase by calling the apply function that we constructed on a vertex using that information. This function was created using the information that we gathered in the phase before this one. The next level is where we are able to put the acquired information to work. In this section of the guide, we will be modifying the values of the vertices by making use of the apply function, which will allow us to do so. There is no actual contact in the parts of the surface that are not joined together by vertices [76].

At this point in the process, we have arrived at the stage that is known as "scattering." During this phase, it is necessary to propagate the modified value (s) to all of the adjacent vertices and edges. This step must be completed before moving on to the next. Some implementations take use of the push-style dispersion, in which updates are pushed to vertices that are located a large distance from one another. This kind of dispersion is advantageous in certain circumstances. Some implementations make use of dispersion in the form described above. As a result of this, some push-style algorithms may choose to skip the collect phase in order to cut down on the total number of edge crossings that are brought about by the situation. It's possible to do this in order to save the available resources [77].

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

Since the GAS model is able to abstract the synchronization cost, it is much simpler to evaluate the complexity and validity of graph algorithms that are implemented in this language. This is the case because the GAS model is able to abstract the synchronization cost. This holds true regardless of whether or not the language itself implements the algorithms in question [76].

However, while developing a method for graph processing that is compatible with this paradigm, we shouldn't ignore the cost of synchronization on GPUs because it has the potential to have a significant impact on the rate at which the implementation is carried out. This is because synchronization on GPUs is required in order for the method to work. This is because synchronization on GPUs is essential in order to guarantee that data is transmitted between threads in a timely way. This is the reason for the aforementioned issue. Sync, for example, should never be used by threads that are outside of the same block since this is the only circumstance in which CUDA advises using it. Instead, it should only be used by threads that are within the same block. The computation is going to be broken up into many different blocks using a number of different kernels, and the machine is going to ensure that all of the threads included inside each of the blocks are consistent with one another as well as the rest of the world [78] .

The termination points of the kernels that are being executed sequentially by the GPU fulfill the purpose of a global barrier as a direct result of this sequential execution of the kernels by the GPU. When used in combination with global synchronization, regional block-wise synchronization leads to a considerable rise in the amount of money spent. Getting rid of the synchronization points that are present in graphs is one of the most essential aims of the desynchronization technique that is being performed [79].

The BSP model is being put into action in a way that ensures it will function correctly by adhering to a predefined set of processes that were developed in the past. Within each superstep, the parallel processes carry out asynchronous execution and communicate with one another and carry out asynchronous execution via messages. Asynchronous execution is carried out by the parallel processes. If a barrier is positioned at the end of each superstep, it is feasible to synchronize all of the operations such that they occur at the same time as one another. In order to be more detailed, each superstep is comprised of the three acts that are described in the following paragraphs [80, 81]:

The process of carrying out computations using hardware that is easily accessible is referred to as "local computing," and this kind of computing is referred to as "local computing."

At this point in the process, communication is taking place on a continental as well as a global scale. Currently. The synchronization of every single calculation and communication in its entirety While we were working on syncing the boundaries throughout this period, At this point, there is not the slightest bit of doubt in anyone's mind that we will be successful in achieving our objectives.

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

3. GRAPH APPLICATION CHALLENGES

The following is a synopsis of the research that has been carried out about the challenges that are connected with the use of GPUs for graph processing:

Because it is constructed using a Single Instruction Multiple Threads architecture, the Graphics Processing Unit (GPU) makes it a great lot easier to distribute work among the numerous cores that are available in a system. This is one of the benefits of using a GPU. This is due to the fact that the GPU was developed using an architecture known as Single Instruction Multiple Threads (SIMTs). When doing computations with graphs using parallel processing, one of the unanticipated results that often occurs is an imbalance in the load. This is due to the fact that graphs have an inherent unbalanced nature. The control unit that is the central processing unit, often known as the CPU for short, is a component that is not only exceptionally powerful but also incredibly versatile. It is possible to make changes to the scheduling strategy whenever such changes are required to be made.

The capability of accessing memory in a certain format the main memory that is incorporated in the most majority of today's central processing units is capable of handling graphs of any size, even those that are quite comprehensive and based on the real world (CPUs). Because of the relatively little amount of memory that a GPU has, doing analysis on really large graphs may be challenging when utilizing one. In order to discover a solution to this issue, it can be helpful to use the strategies for data partitioning that are available. The existing body of scholarly literature offers only a rudimentary solution to the question of how to arrange data that is represented by irregular graphs [82].

In order to provide uninterrupted access to memory, data layout is often handled by means of graph processing. This is done so that the data may be stored in a more organized fashion. In a graphics processing unit (GPU), this component is known as the global shared cache. This cache is used by each and every core that is present. To one's advantage, it is helpful to have a large number of SIMD threads that are able to receive data. Users of graph algorithms and graph data structures have the option to obtain data in a random way when using these two types of graph technology. When it comes to non-standard data structures, the level of parallelism that is necessary in order to take use of the advantages afforded by the GPU is rather modest. This makes it possible to take advantage of the GPU [83, 84].

Miscellaneous. It is possible for the central processing unit to carry out calculations for programs that incorporate branching depending on circumstances since the CPU is flexible enough. It is stated that a branch has diverged when multiple different threads that are contained inside the same condition branch and travelling along the same wave front start going in divergent directions. If this were to be permitted, the performance of the GPU would suffer drastically since it is only able to carry out the processing of a single route at a time when it is working in SIMD mode. If this were to take place, the performance of the GPU would suffer significantly. When parallel GPU programming is being done, the threads are often arranged into blocks to make it easier for them to communicate with one another and share memory. This is done in order to allow the blocks to be moved in various directions. The parameters of

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

the kernel have a considerable impact on the amount of parallelism that may be done, and this effect can be tremendous [85].

4. LITERATURE REVIEW

Research on graphics processing units has been the subject of a significant amount of effort over the course of the last several years (GPUs). In the next paragraphs, we are going to speak about some of the research that has been done, so stay tuned for that!

Z. Zheng and his fellow researchers came up with a high-performance approach for coloring graphs and dubbed it the Feluca technique [86]. This method was named after one of the researchers who worked on the project with them. Combining a recursive strategy with a sequential spread-based process is what this technique does in order to get the outcomes that are desired. This approach uses a spread-based sequential strategy. People will often refer to the approach by its name, which is the Feluca technique. This is because Feluca was the inventor of the system. Feluca starts the process by using a recursive algorithm in order to assign colors to the great majority of the vertices in the network in the very first stage of the process. The procedure will now proceed from this point forward. After then, a method known as sequential spread is used in order to color the vertices in the graph that are still uncolored. This is done so that any inherent conflicts that may come up as a consequence of the recursive process may be avoided.

The presentation of this one-of-a-kind heterogeneous CPU+GPU-enhanced DFTB technique was done with the intention of making the routine and efficient modeling of enormously large chemical and biological systems easier to accomplish. I. Allec & Colleagues. [87]They put a variety of algorithmic implementations, hardware configuration benchmarks, and applications of this technology through its paces utilizing a wide variety of chemical and biological systems in order to do an in-depth analysis of the technology and get an understanding of its value. This was done in order to provide us with a full grasp of and respect for the possibilities afforded by this technology. In conclusion, a large-scale DFTB MD simulation of explicitly resolved HIV protease is offered as a proof-of-concept example of an exceptionally huge and complex system. This simulation has 3,974 atoms in total. This is the first time that every explicitly resolved protein has been explored in detail at the quantum-based MD level, as far as we know, and it is also the first time that this has been done. Because of the combination of a conventional flight computer with the small GPU/SoC systems that are now available on the market, this idea has been realized at the UGA Small Satellite Research Laboratory through the use of integrated GPUs for high-performance computing in LEO. This was made possible as a result of the combination of these two technologies. The use of artificial intelligence in space, the development of computer vision and neural networks, and high-performance computing are just a few of the many tasks that may be accomplished with the assistance of such technology. There are a great number of other potential applications as well. The accomplishment of a considerable number of extra goals is made possible by the use of this form of technological aid. Fluid-implicit particle, more commonly known as FLIP, is a technique that was employed by Wu et al. [13] over a sparse grid topology in order to allow the building of an effective

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

massive-scale fluid simulation making use of GPU technology. This approach is more often referred to as FLIP. In the publication titled NVIDIA GVDB Voxels, this architecture was described in detail. [88] The system is capable of managing tens of millions of particles within a simulation region that is almost unlimited in size, it offers novel approaches to the establishment of a parallel, sparse grid hierarchy, and it offers fast incremental updates for transferring particles to the GPU. These are just a few of the capabilities that the system possesses. These are but a few examples of the capabilities that may be found in the system. These are only a few examples of the features that may be available in the system. There are many more. In addition, as part of the FLIP approach, a matrix-free GPU-based conjugate gradient solver that is tailored for sparse grids has been constructed, which makes it more effective for work. The invention of the FLIP method allowed for the execution of this alteration that was before impossible. The FLIP strategy has been improved to the point that it is now more efficient thanks to these modifications. [Y.-Y. Djenou and his colleagues [89] were successful in efficiently optimizing a colony of bees by relying the majority of the time on the parallelism that was provided by graphics processing units (GPUs). Mining huge datasets was the primary inspiration for the development of the GBSO-Miner, which was created with the purpose of being used in the Earth Observation (EO) market segment of the business. This was done because there has not been a comparable breakthrough in the technology used for on-board data processing and downlink for a considerable amount of time, and as a result, this needed to be done. As a result, this was carried out. In conjunction with parallelized image processing software, a novel GPU-accelerated on-board data processing architecture that was proposed by L. It's Davidson and C. P. Bridges [90]demonstrates that it is possible to achieve both high data processing throughput and high compression efficiency. While [91] shown that it is possible to fulfill both of these goals at the same time, which proves that the validity of this claim was established. This is taken into account by the urban dispersion model, which does this by drawing parallels between the mechanisms of heat transfer and the operation of public transit systems. As a consequence of this, the model is able to take this element into consideration. As a direct result of this, the model is in a position to take into account this component. Throughout the course of their work, Kristóf and B. Papp [92] make use of a piece of software that is known as GNU, which stands for "Graphics Processing Unit." When this piece of software was first developed, it was done so with the intention that it would eventually be used in the area of mechanical engineering. By allowing for the modification of geometry and the observation of transient flow and concentration fields, the simulation software makes it possible to conduct an analysis and assessment of a variety of different approaches to the design of the system that is being modeled. This makes it possible to conduct an analysis and assessment of a variety of different ways to design the system that is being modeled. Because of this, it is now feasible to investigate and assess a large variety of different alternative configurations for the system. We will be making use of the letter in order to determine whether or not the outcomes of the time series include data patterns that are comparable to one another. H. Zhu et al. [93] presented an original approach to the problem that needed to be solved. Create an optimization plan based on the GPU and CUDA, with the ultimate objective of accelerating access to GPU memory serving as the primary focus of your efforts. It is recommended that a GPGPU be used for developing this technique. According to

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

Tsung Tai Yeh et al. [94], Pagoda is a daemon kernel that virtualizes GPU resources. Pagoda was developed by these researchers. Pagoda operates in a way that is comparable to that of an operating system in terms of its functionality. [Tsung Tai Yeh and a number of other people] were responsible for the creation of the Pagoda.

N. Capodieci et al. [95] came up with the name SiGAMMA, which stands for Server-based integrated GPU Arbitration System for Memory Access. SiGAMMA was created by these individuals. Their objective was to eliminate any potential of a conflict arising between the workloads that were being carried out on the CPU and the memory needs that were being imposed by the GPU. Server-based integrated GPU Arbitration System for Memory Access is the abbreviation known as SiGAMMA.

The parallel CCD approach that was developed by [P.] Du et al. [96] has the potential to accelerate the eradication of N-body CCD by dividing the work that needs to be done over a number of high-performance graphics processing units. This would be accomplished by dividing the work that needs to be done. The necessary labor might be divided up amongst many people in order to achieve this goal. This might be accomplished by dividing the work that needs to be done up and assigning it to a number of graphics processing units that have a high overall degree of performance (GPUs).

COMPARISON AND DISCUSSION

- Because of the extensive amount of research that has been done on graphics processing units (GPUs), as well as the fact that GPUs can be utilized in a wide variety of settings, we are only going to go over a select few of these applications in this section. The reason for this is because of the extensive amount of research that has been done on GPUs. Graphics processing units (GPUs) have a wide variety of applications, ranging from video games to medical imaging to artificial intelligence. There are a wide variety of applications that make use of graphics processing units (GPUs), including computer graphics, virtual reality, and medical imaging.
- The coloring of graphs and the processes that take place in the fields of chemistry and biology will both be investigated throughout this lesson. The first topic of discussion will be the coloring of graphs.
- • as an on-board computer that has the potential to be used in far more compact iterations of spacecraft that are now in operation
- The very first system of its sort anywhere in the world, making use of graphics processing units (GPUs) by themselves, and being entirely and completely space-efficient in every conceivable area.
- The following is a summary of some of the things that SiGAMMA may be able to achieve in order to assist in eliminating interference between jobs that are being done by the CPU and the GPU:

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

- The study of how people move through cities; the enhancement of the efficiency with which bee swarms carry out their duties; NASA and other applications of space technology;
- Concerning Particular Occupations, Numerous Alternative Methods for the Analysis of Time-Series Data, and Numerous Other Concerns and Concerning Questions

Table 1: Comparison among related works.

Ref.	field	Used	Result
[86]	Graph Coloring	Feluca Algorithm	improve the graph coloring achieve 1.19-8.39 speedup over the state-of-the-art algorithms.
[87]	large chemical and biological systems	new heterogeneous CPU+GPU	high computational performance
[88]	Space	for Small Satellites as a Flight Computer	to meet many of NASA's goals that require space based AI
[89]	Fast Fluid Simulations	FLIP technique	Faster than running on the CPU.
[90]	bees swarm optimization	GBSO-Miner	up to 800 times faster than an optimized CPU Implementation
[91]	Space application	low power embedded GPU	Reduce data corruption from up to 46% to 2%, execution time overhead of 130%.
[92]	Urban Dispersion Studies	Large Eddy Simulation (LES)	The model results show a satisfactory agreement
[93]	find all similar subsequences in the time series data	Dynamic Time Warping (DTW) algorithm	achieve better performance with using GPU parallelization
[94]	Narrow Tasks	Pagoda	achieves speedup of 5.70x , 1.51x , 1.69x over PThreads, CUDA-HyperQ ,GeMTC Respectively and much lower latency per task
[95]	eliminating the interference on CPU and GPU tasks	SiGAMMA	effective to solve inflation
[96]	Parallel Continuous Collision	parallel CCD algorithm	more computationally efficient than existing sequential CCD approaches.

5. CONCLUSIONS

We give a simplified synopsis of the GPU-related research that has been published across a broad variety of work sectors, including the chemical, biological, space, and other specialized areas of operation, within the constraints of this examination. These disciplines include: This investigation was carried out so that we could provide an answer to the following question:

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

The following types of businesses fall under this category: In this article, a discussion is held on the concepts of graph processing on a graphics processing unit (GPU), as well as its significance, applications, and issues, as well as a paradigm for implementing such processing. The discussion also focuses on a way to implement such processing. The concepts of graph processing on a GPU are the subject of the discussion here. These ideas are discussed in conjunction with a model that may be used to carry out the relevant processing. In addition to that, your attention will be given to a possible answer to a problem that has been offered to you. Someone who is not acquainted with GPU but is interested in learning more about it has a chance of finding the information that has been provided to be helpful. It is possible that this will be the case. The software that is used to administer the cloud not only monitors activities and services, but it also creates dynamic configurations of the cloud, which is something that is very necessary in order to increase the efficiency of operational procedures. In other words, the software that is used to administer the cloud creates dynamic configurations in addition to monitoring activities and services. This is made possible via the use of virtual machines, more often referred to as VMs, which are able to save data and carry out processing operations on the data while being in the cloud. In this article, we discussed the potential repercussions that cloud-ready software and tools may have on the operation of a business, as well as the potential benefits that may be obtained by making use of cloud-ready software and technologies. In addition, we also discussed the potential advantages that may be obtained by making use of cloud-ready software and technologies. In addition, we highlighted the possible benefits that may be realized by making use of software and technologies that are cloud-ready.

References

[1] F. E. F. Samann, S. R. Zeebaree, and S. Askar, "IoT provisioning QoS based on cloud and fog computing," Journal of Applied Science and Technology Trends, vol. 2, pp. 29-40, 2021.

[2] K. Jacksi, N. Dimililer, and S. Zeebaree, "State of the art exploration systems for linked data: a review," Int. J. Adv. Comput. Sci. Appl. IJACSA, vol. 7, pp. 155-164, 2016.

[3] R. Ibrahim, S. Zeebaree, and K. Jacksi, "Survey on semantic similarity based on document clustering," Adv. sci. technol. eng. syst. j, vol. 4, pp. 115-122, 2019.

[4] S. A. Mohammed, R. H. Saeed, J. A. Ahmed, S. B. Muhammad, Z. S. Ageed, and Z. N. Rashid, "GPU Concepts and Graph Application Challenges: A."

[5] N. R. Omar, R. H. Saeed, J. A. Ahmed, S. B. Muhammad, Z. S. Ageed, and Z. N. Rashid, "Enhancing OS Memory Management Performance: A."

[6] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, "Enhance the mammogram images for both segmentation and feature extraction using wavelet transform," in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 100-105.

[7] H. S. Yahia, S. Zeebaree, M. Sadeeq, N. Salim, S. F. Kak, A. Adel, et al., "Comprehensive survey for cloud computing based nature-inspired algorithms optimization scheduling," Asian Journal of Research in Computer Science, vol. 8, pp. 1-16, 2021.

[8] H. M. Yasin, S. Zeebaree, M. Sadeeq, S. Y. Ameen, I. M. Ibrahim, R. R. Zebari, et al., "IoT and ICT based smart water management, monitoring and controlling system: A review," Asian Journal of Research in Computer Science, vol. 8, pp. 42-56, 2021.

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

[9] L. M. Abdulrahman, S. R. Zeebaree, S. F. Kak, M. A. Sadeeq, A. AL-Zebari, B. W. Salim, et al., "A state of art for smart gateways issues and modification," Asian Journal of Research in Computer Science, vol. 7, pp. 1-13, 2021.

[10] S. Zeebaree, S. Ameen, and M. Sadeeq, "Social media networks security threats, risks and recommendation: A case study in the Kurdistan region," International Journal of Innovation, Creativity and Change, vol. 13, pp. 349-365, 2020.

[11] L. M. Haji, O. M. Ahmad, S. Zeebaree, H. I. Dino, R. R. Zebari, and H. M. Shukur, "Impact of cloud computing and internet of things on the future internet," Technology Reports of Kansai University, vol. 62, pp. 2179-2190, 2020.

[12] F. Q. Kareem, S. Zeebaree, H. I. Dino, M. Sadeeq, Z. N. Rashid, D. A. Hasan, et al., "A survey of optical fiber communications: challenges and processing time influences," Asian Journal of Research in Computer Science, pp. 48-58, 2021.

[13] R. J. Hassan, S. R. Zeebaree, S. Y. Ameen, S. F. Kak, M. A. Sadeeq, Z. S. Ageed, et al., "State of art survey for iot effects on smart city technology: challenges, opportunities, and solutions," Asian Journal of Research in Computer Science, vol. 8, pp. 32-48, 2021.

[14] K. Jacksi, S. R. Zeebaree, and N. Dimililer, "LOD Explorer: Presenting the Web of Data," Int. J. Adv. Comput. Sci. Appl. IJACSA, vol. 9, pp. 1-7, 2018.

[15] A. Zeebaree, A. Adel, K. Jacksi, and A. Selamat, "Designing an ontology of E-learning system for duhok polytechnic university using protégé OWL tool," J Adv Res Dyn Control Syst Vol, vol. 11, pp. 24-37, 2019.

[16] H. M. Yasin, S. R. Zeebaree, and I. M. Zebari, "Arduino based automatic irrigation system: Monitoring and SMS controlling," in 2019 4th Scientific International Conference Najaf (SICN), 2019, pp. 109-114.

[17] A. AL-Zebari, S. Zeebaree, K. Jacksi, and A. Selamat, "ELMS–DPU ontology visualization with Protégé VOWL and Web VOWL," Journal of Advanced Research in Dynamic and Control Systems, vol. 11, pp. 478-85, 2019.

[18] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, "Multi-level of DNA encryption technique based on DNA arithmetic and biological operations," in 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 312-317.

[19] S. Zebari and N. O. Yaseen, "Effects of parallel processing implementation on balanced load-division depending on distributed memory systems," J. Univ. Anbar Pure Sci, vol. 5, pp. 50-56, 2011.

[20] K. Jacksi, N. Dimililer, and S. R. Zeebaree, "A survey of exploratory search systems based on LOD resources," 2015.

[21] K. B. Obaid, S. Zeebaree, and O. M. Ahmed, "Deep learning models based on image classification: a review," International Journal of Science and Business, vol. 4, pp. 75-81, 2020.

[22] O. H. Jader, S. Zeebaree, and R. R. Zebari, "A state of art survey for web server performance measurement and load balancing mechanisms," International Journal of Scientific & Technology Research, vol. 8, pp. 535-543, 2019.

[23] M. A. Omer, S. R. Zeebaree, M. A. Sadeeq, B. W. Salim, Z. N. Rashid, and L. M. Haji, "Efficiency of malware detection in android system: A survey," Asian Journal of Research in Computer Science, vol. 7, pp. 59-69, 2021.

[24] Z. S. Ageed, S. Zeebaree, M. Sadeeq, M. B. Abdulrazzaq, B. W. Salim, A. A. Salih, et al., "A state of art survey for intelligent energy monitoring systems," Asian Journal of Research in Computer Science, vol. 8, pp. 46-61, 2021.

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

[25] S. Zeebaree, R. R. Zebari, K. Jacksi, and D. A. Hasan, "Security Approaches For Integrated Enterprise Systems Performance: A Review," Int. J. Sci. Technol. Res, vol. 8, 2019.

[26] K. Jacksi, R. K. Ibrahim, S. R. Zeebaree, R. R. Zebari, and M. A. Sadeeq, "Clustering documents based on semantic similarity using HAC and K-mean algorithms," in 2020 International Conference on Advanced Science and Engineering (ICOASE), 2020, pp. 205-210.

[27] P. Y. Abdullah, S. Zeebaree, H. M. Shukur, and K. Jacksi, "HRM system using cloud computing for Small and Medium Enterprises (SMEs)," Technology Reports of Kansai University, vol. 62, p. 04, 2020.

[28] S. R. Zeebaree, A. B. Sallow, B. K. Hussan, and S. M. Ali, "Design and simulation of high-speed parallel/sequential simplified DES code breaking based on FPGA," in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 76-81.

[29] Z. S. AGEED, Z. N. RASHID, R. T. HASAN, Y. S. JGHEF, T. M. G. SAMI, and M. A. OMER, "WEB TECHNOLOGY DEVELOPMENT AND DATA-INTENSIVE INFLUENCES ON SEMANTIC INFORMATION RETRIEVING BASED ON PARALLEL AND CLOUD COMPUTING."

[30] S. R. Zeebaree, H. M. Shukur, L. M. Haji, R. R. Zebari, K. Jacksi, and S. M. Abas, "Characteristics and analysis of hadoop distributed systems," Technology Reports of Kansai University, vol. 62, pp. 1555-1564, 2020.

[31] Z. S. Ageed, S. R. Zeebaree, M. A. Sadeeq, R. K. Ibrahim, H. M. Shukur, and A. Alkhayyat, "Comprehensive Study of Moving from Grid and Cloud Computing Through Fog and Edge Computing towards Dew Computing," in 2021 4th International Iraqi Conference on Engineering Technology and Their Applications (IICETA), 2021, pp. 68-74.

[32] G. M. Zebari, S. Zeebaree, M. M. Sadeeq, and R. Zebari, "Predicting Football Outcomes by Using Poisson Model: Applied to Spanish Primera División," Journal of Applied Science and Technology Trends, vol. 2, pp. 105-112, 2021.

[33] S. H. Ahmed and S. Zeebaree, "A survey on security and privacy challenges in smarthome based IoT," International Journal of Contemporary Architecture, vol. 8, pp. 489-510, 2021.

[34] Z. S. Ageed, S. R. Zeebaree, M. M. Sadeeq, S. F. Kak, H. S. Yahia, M. R. Mahmood, et al., "Comprehensive survey of big data mining approaches in cloud systems," Qubahan Academic Journal, vol. 1, pp. 29-38, 2021.

[35] A. I. Abdulla, A. S. Abdulraheem, A. A. Salih, M. Sadeeq, A. J. Ahmed, B. M. Ferzor, et al., "Internet of things and smart home security," Technol. Rep. Kansai Univ, vol. 62, pp. 2465-2476, 2020.

[36] N. T. Muhammed, S. R. Zeebaree, and Z. N. Rashid, "Distributed Cloud Computing and Mobile Cloud Computing: A Review," QALAAI ZANIST JOURNAL, vol. 7, pp. 1183-1201, 2022.

[37] Y. S. Jghef, M. J. M. Jasim, H. M. Ghanimi, A. D. Algarni, N. F. Soliman, W. El-Shafai, et al., "Bio-Inspired Dynamic Trust and Congestion-Aware Zone-Based Secured Internet of Drone Things (SIoDT)," Drones, vol. 6, p. 337, 2022.

[38] S. Chavhan, S. R. Zeebaree, A. Alkhayyat, and S. Kumar, "Design of Space Efficient Electric Vehicle Charging Infrastructure Integration Impact on Power Grid Network," Mathematics, vol. 10, p. 3450, 2022.

[39] L. M. Abdulrahman, S. R. Zeebaree, and N. Omar, "State of Art Survey for Designing and Implementing Regional Tourism Web based Systems," Academic Journal of Nawroz University, vol. 11, pp. 100-112, 2022.

[40] A. S. Aljuboury, S. R. Zeebaree, F. Abedi, Z. S. Hashim, R. Q. Malik, I. K. Ibraheem, et al., "A New Nonlinear Controller Design for a TCP/AQM Network Based on Modified Active Disturbance Rejection Control," Complexity, vol. 2022, 2022.

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

[41] R. K. Ibrahim, S. R. Zeebaree, K. Jacksi, S. H. Ahmed, S. M. Mohammed, R. R. Zebari, et al., "Clustering Document based on Semantic Similarity Using Graph Base Spectral Algorithm," in 2022 5th International Conference on Engineering Technology and its Applications (IICETA), 2022, pp. 254-259.

[42] H. B. Abdalla, A. M. Ahmed, S. R. Zeebaree, A. Alkhayyat, and B. Ihnaini, "Rider weed deep residual network-based incremental model for text classification using multidimensional features and MapReduce," PeerJ Computer Science, vol. 8, p. e937, 2022.

[43] K. P. M. Kumar, J. Mahilraj, D. Swathi, R. Rajavarman, S. R. Zeebaree, R. R. Zebari, et al., "Privacy Preserving Blockchain with Optimal Deep Learning Model for Smart Cities," CMC-COMPUTERS MATERIALS & CONTINUA, vol. 73, pp. 5299-5314, 2022.

[44] V. D. Majety, N. Sharmili, C. R. Pattanaik, E. L. Lydia, S. R. Zeebaree, S. N. Mahmood, et al., "Ensemble of Handcrafted and Deep Learning Model for Histopathological Image Classification," CMC-COMPUTERS MATERIALS & CONTINUA, vol. 73, pp. 4393-4406, 2022.

[45] D. M. Abdulqader and S. R. Zeebaree, "Impact of Distributed-Memory Parallel Processing Approach on Performance Enhancing of Multicomputer-Multicore Systems: A Review," QALAAI ZANIST JOURNAL, vol. 6, pp. 1137-1140, 2021.

[46] H. Abdullah and S. R. Zeebaree, "Android Mobile Applications Vulnerabilities and Prevention Methods: A Review," 2021 2nd Information Technology To Enhance e-learning and Other Application (IT-ELA), pp. 148-153, 2021.

[47] Z. N. Rashid, S. R. Zeebaree, M. A. Sadeeq, R. R. Zebari, H. M. Shukur, and A. Alkhayyat, "Cloud-based Parallel Computing System Via Single-Client Multi-Hash Single-Server Multi-Thread," in 2021 International Conference on Advance of Sustainable Engineering and its Application (ICASEA), 2021, pp. 59-64.

[48] O. U. Oluoha, T. S. Yange, G. E. Okereke, and F. S. Bakpo, "Cutting Edge Trends in Deception Based Intrusion Detection Systems—A Survey," Journal of Information Security, vol. 12, pp. 250-269, 2021.

[49] S. Zeebaree, N. Cavus, and D. Zebari, "Digital Logic Circuits Reduction: A Binary Decision Diagram Based Approach," LAP LAMBERT Academic Publishing, 2016.

[50] A. M. Abed, Z. N. Rashid, F. Abedi, S. R. Zeebaree, M. A. Sahib, A. J. a. Mohamad Jawad, et al., "Trajectory tracking of differential drive mobile robots using fractional-order proportional-integral-derivative controller design tuned by an enhanced fruit fly optimization," Measurement and Control, vol. 55, pp. 209-226, 2022.

[51] M. Shallal, "2018 International Conference on Advance of Sustainable Engineering and its Application (ICASEA)," 2018.

[52] Z. N. Rashid, S. Zeebaree, and A. Sengur, "Novel remote parallel processing code-breaker system via cloud computing," Ed: TRKU, 2020.

[53] H. Fawzia, D. Ahmeda, S. A. Mostafac, M. F. M. Fudzeec, M. A. Mahmoodd, S. R. Zeebaree, et al., "A REVIEW OF AUTOMATED DECISION SUPPORT TECHNIQUES FOR IMPROVING TILLAGE OPERATIONS," REVISTA AUS, vol. 26, pp. 219-240, 2019.

[54] N. M. ABDULKAREEM and S. R. ZEEBAREE, "OPTIMIZATION OF LOAD BALANCING ALGORITHMS TO DEAL WITH DDOS ATTACKS USING WHALE OPTIMIZATION ALGORITHM," Journal of Duhok University, vol. 25, pp. 65-85, 2022.

[55] A. A. Yazdeen and S. R. Zeebaree, "Comprehensive Survey for Designing and Implementing Web-based Tourist Resorts and Places Management Systems," Academic Journal of Nawroz University, vol. 11, pp. 113-132, 2022.

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

[56] R. R. Zebari, S. R. Zeebaree, Z. N. Rashid, H. M. Shukur, A. Alkhayyat, and M. A. Sadeeq, "A Review on Automation Artificial Neural Networks based on Evolutionary Algorithms," in 2021 14th International Conference on Developments in eSystems Engineering (DeSE), 2021, pp. 235-240.

[57] S. I. Ahmed, S. Y. Ameen, and S. R. Zeebaree, "5G Mobile Communication System Performance Improvement with Caching: A Review," in 2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI), 2021, pp. 1-8.

[58] H. A. Mohammed, S. Zeebaree, V. M. Tiryaki, and M. M. Sadeeq, "Web-Based Land Registration Management System: Iraq/Duhok Case Study," Journal of Applied Science and Technology Trends, vol. 2, pp. 113-119, 2021.

[59] R. K. Ibrahim, S. R. Zeebaree, K. Jacksi, M. A. Sadeeq, H. M. Shukur, and A. Alkhayyat, "Clustering document based semantic similarity system using TFIDF and k-mean," in 2021 International Conference on Advanced Computer Applications (ACA), 2021, pp. 28-33.

[60] O. H. Jader, S. R. Zeebaree, R. R. Zebari, H. M. Shukur, Z. N. Rashid, M. A. Sadeeq, et al., "Ultra-Dense Request Impact on Cluster-Based Web Server Performance," in 2021 4th International Iraqi Conference on Engineering Technology and Their Applications (IICETA), 2021, pp. 252-257.

[61] A. E. Mehyadin, S. R. Zeebaree, M. A. Sadeeq, H. M. Shukur, A. Alkhayyat, and K. H. Sharif, "State of Art Survey for Deep Learning Effects on Semantic Web Performance," in 2021 7th International Conference on Contemporary Information Technology and Mathematics (ICCITM), 2021, pp. 93-99.

[62] M. A. Sadeeq and S. R. Zeebaree, "Design and analysis of intelligent energy management system based on multi-agent and distributed iot: Dpu case study," in 2021 7th International Conference on Contemporary Information Technology and Mathematics (ICCITM), 2021, pp. 48-53.

[63] I. M. Ibrahim, S. R. Zeebaree, H. M. Yasin, M. A. Sadeeq, H. M. Shukur, and A. Alkhayyat, "Hybrid Client/Server Peer to Peer Multitier Video Streaming," in 2021 International Conference on Advanced Computer Applications (ACA), 2021, pp. 84-89.

[64] A. R. Brodtkorb, T. R. Hagen, C. Schulz, and G. Hasle, "GPU computing in discrete optimization. Part I: Introduction to the GPU," EURO journal on transportation and logistics, vol. 2, pp. 129-157, 2013.

[65] J. D. Owens, M. Houston, D. Luebke, S. Green, J. E. Stone, and J. C. Phillips, "GPU computing," Proceedings of the IEEE, vol. 96, pp. 879-899, 2008.

[66] D. M. Abdullah and S. R. Zeebaree, "Comprehensive survey of IoT based arduino applications in healthcare monitoring."

[67] C. Hong, D. Chen, W. Chen, W. Zheng, and H. Lin, "MapCG: Writing parallel program portable between CPU and GPU," in Proceedings of the 19th international conference on Parallel architectures and compilation techniques, 2010, pp. 217-226.

[68] J. Choi, Z. Fink, S. White, N. Bhat, D. F. Richards, and L. V. Kale, "Gpu-aware communication with ucx in parallel programming models: Charm++, mpi, and python," in 2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 2021, pp. 479-488.

[69] B. W. SALIM and S. R. ZEEBAREE, "ISOLATED AND CONTINUOUS HAND GESTURE RECOGNITION BASED ON DEEP LEARNING: A REVIEW."

[70] F. Abedi, S. R. Zeebaree, Z. S. Ageed, H. M. Ghanimi, A. Alkhayyat, M. A. Sadeeq, et al., "Severity Based Light-Weight Encryption Model for Secure Medical Information System."

[71] Y. Sun, T. Baruah, S. A. Mojumder, S. Dong, X. Gong, S. Treadway, et al., "MGPUSim: enabling multi-GPU performance modeling and optimization," in Proceedings of the 46th International Symposium on Computer Architecture, 2019, pp. 197-209.

The Seybold Report
ISSN 1533-9211
DOI 10.17605/OSF.IO/7E5Q9

[72] J. Kim, S. Seo, J. Lee, J. Nah, G. Jo, and J. Lee, "OpenCL as a unified programming model for heterogeneous CPU/GPU clusters," ACM SIGPLAN Notices, vol. 47, pp. 299-300, 2012.

[73] Z. S. Hammed, S. Y. Ameen, and S. R. Zeebaree, "Investigation of 5G Wireless Communication with Dust and Sand Storms," Journal of Communications, vol. 18, 2023.

[74] I. S. Abdulkhaleq and S. R. Zeebaree, "Science and Business," International Journal, vol. 5, pp. 126-136.

[75] Y. Wu, Y. Wang, Y. Pan, C. Yang, and J. D. Owens, "Performance characterization of high-level programming models for GPU graph analytics," in 2015 IEEE International Symposium on Workload Characterization, 2015, pp. 66-75.

[76] N. A. Kako, "DDLS: Distributed Deep Learning Systems: A Review," Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol. 12, pp. 7395-7407, 2021.

[77] M. Mehrtash, A. Kargarian, and A. J. Conejo, "Graph-based second-order cone programming model for resilient feeder routing using GIS data," IEEE Transactions on Power Delivery, vol. 35, pp. 1999-2010, 2019.

[78] K. Ekanadham, W. P. Horn, M. Kumar, J. Jann, J. Moreira, P. Pattnaik, et al., "Graph programming interface (gpi) a linear algebra programming model for large scale graph computations," in Proceedings of the ACM International Conference on Computing Frontiers, 2016, pp. 72-81.

[79] H. A. Hussein, S. R. Zeebaree, M. A. Sadeeq, H. M. Shukur, A. Alkhayyat, and K. H. Sharif, "An investigation on neural spike sorting algorithms," in 2021 International Conference on Communication & Information Technology (ICICT), 2021, pp. 202-207.

[80] A. Mohan and G. Remya, "A review on large scale graph processing using big data based parallel programming models," International Journal of Intelligent Systems and Applications, vol. 9, p. 49, 2017.

[81] A. F. Jahwar and S. Zeebaree, "A state of the art survey of machine learning algorithms for IoT security," Asian J. Res. Comput. Sci, pp. 12-34, 2021.

[82] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," Proceedings of the IEEE, vol. 106, pp. 808-828, 2018.

[83] P. Pradhyumna and G. Shreya, "Graph neural network (GNN) in image and video understanding using deep learning for computer vision applications," in 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021, pp. 1183-1189.

[84] F. M. Al-Fiky and Z. S. Ageed, "AUTHENTICATION SYSTEM DEPEND ON HAND GEOMETRY USING BACK PROPAGATION NEURAL NETWORK," 2012.

[85] S. Sahu, A. Mhedhbi, S. Salihoglu, J. Lin, and M. T. Özsu, "The ubiquity of large graphs and surprising challenges of graph processing: extended survey," The VLDB journal, vol. 29, pp. 595-618, 2020.

[86] Z. Zheng, X. Shi, L. He, H. Jin, S. Wei, H. Dai, et al., "Feluca: A two-stage graph coloring algorithm with color-centric paradigm on gpu," IEEE Transactions on Parallel and Distributed Systems, vol. 32, pp. 160-173, 2020.

[87] S. I. Allec, Y. Sun, J. Sun, C.-e. A. Chang, and B. M. Wong, "Heterogeneous CPU+ GPU-enabled simulations for DFTB molecular dynamics of large chemical and biological systems," Journal of chemical theory and computation, vol. 15, pp. 2807-2815, 2019.

[88] C. Adams, A. Spain, J. Parker, M. Hevert, J. Roach, and D. Cotten, "Towards an integrated GPU accelerated SoC as a flight computer for small satellites," in 2019 IEEE Aerospace Conference, 2019, pp. 1-7.

[89] K. Wu, N. Truong, C. Yuksel, and R. Hoetzlein, "Fast fluid simulations with sparse volumes on the GPU," in Computer Graphics Forum, 2018, pp. 157-167.

ISSN 1533-9211

DOI 10.17605/OSF.IO/7E5Q9

[90] Y. Djenouri, D. Djenouri, A. Belhadi, P. Fournier-Viger, J. C.-W. Lin, and A. Bendjoudi, "Exploiting GPU parallelism in improving bees swarm optimization for mining big transactional databases," Information Sciences, vol. 496, pp. 326-342, 2019.

[91] R. L. Davidson and C. P. Bridges, "Error resilient GPU accelerated image processing for space applications," IEEE Transactions on Parallel and Distributed Systems, vol. 29, pp. 1990-2003, 2018.

[92] G. Kristóf and B. Papp, "Application of GPU-based large eddy simulation in urban dispersion studies," Atmosphere, vol. 9, p. 442, 2018.

[93] H. Zhu, Z. GU, H. Zhao, K. Chen, C.-T. Li, and L. He, "Developing a pattern discovery method in time series data and its GPU acceleration," Big Data mining and Analytics, vol. 1, pp. 266-283, 2018.

[94] X. Tian, Y. Huang, and H. Wang, "Deviation of Chinese adults' diet from the Chinese food pagoda 2016 and its association with adiposity," Nutrients, vol. 9, p. 995, 2017.

[95] N. Capodieci, R. Cavicchioli, P. Valente, and M. Bertogna, "SiGAMMA: Server based integrated GPU arbitration mechanism for memory accesses," in Proceedings of the 25th International Conference on Real-Time Networks and Systems, 2017, pp. 48-57.

[96] P. Du, E. S. Liu, and T. Suzumura, "Parallel continuous collision detection for high-performance GPU cluster," in Proceedings of the 21st ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, 2017, pp. 1-7.