PERFORMANCE ANALYSIS AND BENCH MARKING OF THE MICRO VIRTUAL MACHINES BY PUBLIC CLOUD SERVICE PROVIDERS

A
DISSERTATION

Submitted in partial fulfillment of requirement for the award of Degree of
Master of Technology
In
Department of Computer Science & Engineering

Submitted to

Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, M.P.

Submitted by
Mr. YUGANSH GARG
Enrollment No. – 0187CS20MT06

Under the Supervision of
MS BHAVANA GUPTA
ASSISTANT PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SAGAR INSTITUTE OF SCIENCE & TECHNOLOGY GANDHI NAGAR, BHOPAL, M.P.

Jan 2023

ABSTRACT

Modern IT industries are highly dependent on cloud services including Infrastructure-as-a-Service (IaaS), computing resources such as CPU, processing time, disk space or network. These capabilities can be acquired and released as self-service mostly through application programming interfaces (API) in the form of virtual machines (VMs). Virtual machines are typically available in different configurations or sizes, which are also known as instance types, machine types, or variants. This variety moves from small VMs with less than 1 (shared) CPU core and 1 GB RAM (eg f1-micro) to super large VMs with 128 CPU cores and 1952 GB RAM (eg x1.32xlarge). Configuring a virtual machine for an application is a non-trivial challenge.

The performance of virtual machines may affect the services offered. The Research work attempts to analyze the performance of Light virtual machines provided by different cloud vendors and maps their ability to handle a certain set of tasks. The main goal is to compare the basic virtual machines that are mostly used across the industry for different purposes and find the one that is efficient and productive. The system tested are machines with 1 GB of RAM provided by the most popular cloud service providers.

The goal is to compare identify the virtual machines in this segment which meets the criteria like faster processing, better performance, ease of access and affordable pricing. Thus the machines are put under the rigorous test and the performance is measured and analyzed.

Keywords: *Cloud Computing, Virtualization, IaaS, Public Clouds, Virtual Machines, Benchmarking, Performance Analysis*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

VMs – Virtual Machines

AWS – Amazon Web Services

GCP – Google Cloud Platforms

EC2 – Elastic Cloud Computing

GCE – Google Compute Engine

VMM – Virtual Machine Manager

IaaS – Infrastructure as a Service

CHAPTER 1INTRODUCTION

1.1 Overview

The past two decades have seen increasing number of the online services. The Internet has changed from a mere useful tool to a dominant part of life and culture. To support this phenomenal growth, the handful of computers that were once used to serve the entire online community has turned into clouds spanning hundreds of thousands of servers in stadium-sized data centers. Servers run around the clock, processing millions of requests and providing access to peta bytes of data to users around the world. Driven by the ever-increasing demand for more servers and larger data centres, academia, and industry are directing significant efforts to improve the state-of-the-art server hardware architecture and software design.

The Thesis tries to compare the entry level virtual machines with the configuration 1 core, 1 Gb RAM and minimal storage along with provided network bandwidth. The Virtual Machines provided by the different cloud vendors comes with the different set of attributes features etc.

1.2 Motivation

Clouds are the most important part of the modern IT infrastructure and almost all sectors of modern economies are depended on clouds provide and manage their services. Thus, the performance of the clouds can impact a business in particular and society as whole.

Majority of the small business, start ups use small virtual machines from public cloud vendors. Thus, machines of this segment play a decisive role in the small business. However, lack of exploration in performance analysis and Benchmarking of the micro and small categories of VMs motivated me to Benchmark and analyse the performance of the virtual machines of this segment.

1.3 Thesis Objective

The objective of the thesis is to evaluate the performance of the Virtual machines in the micro and small category. Meaning there by the thesis is directed towards analysis of V.Ms having smaller configuration of RAM 1 GB and single or dual Virtual Core

and storage up to 8-10 GBs. The objective is comparing the performance of the Virtual machines provided by Vendors including Amazon Web Services, MS Azure, Google Cloud Platform, IBM Cloud and Oracle Cloud Platform.

The objective is to fill the research gap in exploration, Performance analysis and Benchmarking in the field of Cloud Bench marking.

1.4 Cloud Computing

Cloud computing is a version for allowing ubiquitous, convenient, on-call for community get entry to a shared pool of configurable computing resources (e.g., networks, servers, garage, packages, and offerings) that can be unexpectedly provisioned and launched with minimum control attempt or carrier issuer interplay. This cloud model is composed of five essential traits, 3 service fashions, and four deployment fashions. The details of cloud computing are explained in chapter 2. [1]

1.5 Performance Analysis

Performance analysis is a method to compare a Public Cloud process based on performance rather than the number of services they provide, the goal is to compare a machine's ability to handle multiple processes simultaneously. Comparing cloud providers and the cloud computing services they offer can often be difficult. Procurement professionals have the expertise to conduct vendor due diligence necessary to determine the health, reliability, and overall capabilities of a vendor's business and pricing, but cloud services have There are a number of fundamental aspects that need attention.[2]

Cloud performance monitoring and testing tools help organizations gain insight into their cloud environments using specific metrics and techniques to evaluate performance. Effective cloud performance is critical to maintaining business continuity and ensuring that all relevant parties have access to cloud services. This applies to basic cloud use of public clouds and complex hybrid cloud and multi-cloud architectures.

Cloud performance metrics allow user to effectively monitor the cloud resources to ensure smooth communication between all components. Cloud performance metrics typically measure input/output operations per second (IOPS), file system performance, caching, and autoscaling.[3]

Detailed performance analysis is a complex process, but the goal is very simple: to identify the root cause of the problem. To do this, we proceed in an orderly step-by-step manner to the home domain of the fault containing the root cause. Since this chapter is specifically about virtualization and cloud computing, we will discuss how these two domains affect this process. But first, let's look at the root cause analysis flow at a high level.[4]

1.6 Benchmarking

Comparing cloud services is the way to systematically study the quality of cloud services based on experiments. For this the benchmark tool creates an artificial load on the cloud service under test - or the System Under Test (SUT) - while being careful tracking detailed quality metrics A key goal of cloud service design benchmarking consists of imitating an application as closely as possible in order to obtain a meaningful results, however benchmark runs are also intended to greatly emphasize service, for example through system load or even injected faults.[5]

When we look at raw computing power, we're talking about CPU and RAM. Performance differences at a purely computational level between clouds are not really that big. However, there are some factors that make real differences.

CHAPTER 2
CLOUDS, VIRTUALIZATION AND IAAS

2.1 Cloud Computing

In cloud computing, user can develop new packages and services save and backup data, host blogs, and websites, examine the statistics, and stream audio and video. There's no want for huge and complicated cloud computing infrastructure as the storage and processing are accomplished on far flung servers referred as cloud servers.

The software program platform and databases are hosted remotely, the person computers are free from the garage and computing responsibilities. Cloud computing offers a huge range of advantages, including high efficiency, price discount, information security, scalability, mobility, catastrophe recovery, and whole manipulate over the data.

Compared to traditional IT, cloud computing provides the organizations a number of advantages: the pocket friendly of paying only for the resources and services used to bring critical applications and services to market faster. the possibility of easy, affordable and - with the right cloud provider - globally; and much more. To offer sensitive/critical apps or data to clients, business partners, or workers, many organisations find extra advantages in combining public cloud services obtained from a cloud service provider with private cloud infrastructure they manage themselves[6]

2.1.1 Cloud Deployment Models

Depending on the method of deployment and access the clouds can be categories into three main Categories Public Clouds, Private Cloud, Hybrid Cloud and community Cloud.

2.1.1.1 Public Cloud

A public cloud is a type of cloud computing in which a cloud service provider makes computing resources available to users – anything from SaaS applications to individual virtual machines (VMs), bare metal computing hardware, to complete enterprise infrastructures and development platforms. Over the public internet

network. These resources may be accessible for free, or access may be sold under subscription-based or pay-per-use pricing models.

A public cloud provider owns, manages, and takes full responsibility for the data centers, hardware, and infrastructure that run its customers' workloads, and typically provides broadband network connectivity to ensure high performance and fast access to applications and data.

Public cloud is a multi-tenant environment –the cloud provider's data center infrastructure is shared by all public/users cloud customers. In the leading public clouds—Amazon Web Services (AWS), Google Cloud, IBM Cloud, Microsoft Azure, and Oracle Cloud—these customers can number in the millions.[1], [7]

2.1.1.2 Private Clouds

A private cloud is a cloud environment in which all cloud infrastructure and computing resources are reserved and accessible only to one customer. A private cloud combines many of the benefits of cloud computing—including flexibility, scalability, and ease of provisioning—with the access control, security, and customization of on-premises infrastructure resources.[1],

A private cloud is typically hosted locally in a customer's data center. However, a private cloud can also be hosted on the infrastructure of an independent cloud provider or built on leased infrastructure located in an off-site data center.

Many companies choose private cloud over public cloud because private cloud is an easier way (or the only way) to meet their compliance requirements. Others choose a private cloud because their workloads deal with confidential documents, biometrics, medical records, financial data, or other sensitive data.

By building a private cloud architecture based on cloud-native principles, an organization gives itself the flexibility to easily move workloads to the public cloud or run them in a hybrid cloud environment whenever they are ready. [1], [6]

2.1.1.3 Hybrid Cloud

A hybrid cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities but are connected by standardized or proprietary technology that enables data and application portability (eg, clouds).[1]

It combines the best cloud services and features from multiple cloud computing vendors choose the optimal cloud computing environment for each workload and freely move workloads between public and private clouds as circumstances change. Hybrid cloud and hybrid multicloud in particular – helps a company achieve its technical and business goals more efficiently and cost-effectively than public cloud or private cloud alone [9]

2.1.1.4 Community Cloud

A cloud infrastructure is established for the exclusive use of a specific community of consumers from organizations that share interests (eg, mission, security requirements, policies and compliance aspects). It may be owned, managed and operated by one or more organizations in the community, a third party or some combination thereof and may exist on or off premises. [1]

Service providers can target new client segments that could not be reached with a standard cloud offering. Although the investment in the underlying infrastructure and security processes of the cloud may be higher, the competition in this segment is less and the market prices are potentially higher than for standard public cloud services. Depending on the target industry, the cloud services offered are evolving from a commodity business to a high-value, high-margin business that may be more attractive to service providers. [10]

2.1.2 Cloud Service Models

2.1.2.1 Infrastructure as a Service (IaaS)

Infrastructure-as-a-Service, commonly referred to simply as "IaaS," is a form of cloud computing that provides consumers with basic computing, networking, and storage resources on demand, over the Internet, and for a fee. You-go base. IaaS allows end-users to scale up and down resources as needed, reducing the need for high initial capital expenditures or unnecessary "proprietary" infrastructure, especially for "peak" workloads. Unlike PaaS and SaaS (even newer computing models such as containers and serverless models), IaaS provides the lowest level of resource management in the cloud. .[1][11])

The option provided to the consumer is to use the provider's applications running on the cloud infrastructure. Applications are accessed from various client devices either through a thin client interface such as a web browser (eg web email) or a program

interface. The consumer does not manage or control the underlying cloud infrastructure, including the network, servers, operating systems, storage, or even individual application functions, with the possible exception of limited user application configuration settings. [1]

The main advantage of SaaS is that it transfers all infrastructure and application management to the SaaS vendor. All the user has to do is create an account, pay the fee and start using the app. The vendor handles everything else, from maintaining server hardware and software to managing user access and security, storing and managing data, implementing upgrades and patches, and more.[12]

2.2 Infrastructure as a Service (IaaS)

NIST defines “IaaS as a consumer-provided capability to provide processing, storage, networking, and other underlying computing resources where the consumer is able to deploy and run any software, which may include operating systems and applications”. The customer does not manage or control the underlying cloud infrastructure, but has control over operating systems, storage and deployed applications; and possibly limited control of selected network components (eg host firewalls).[1]

Infrastructure-as-a-Service, typically mentioned without a doubt as "IaaS," is a form of cloud computing that offers clients with basic computing, networking, and garage resources on demand, over the Internet, and for a price. You-cross base. IaaS permits end customers to scale up and down resources as wished, lowering the need for excessive initial capital expenditure or needless "proprietary" infrastructure, especially for "top" workloads. Unlike PaaS and SaaS (or even newer computing models including boxes and serverless models), IaaS presents the lowest stage of cloud useful resource management. [13]

IaaS clients can normally select between digital machines (VMs) hosted on shared physical hardware (the cloud carrier provider manages the virtualization) or bare servers on dedicated (non-shared) physical hardware. Customers can provision, configure, and perform servers and infrastructure assets via a graphical dashboard or programmatically via application programming interfaces (APIs).[21].

IaaS can be thought of as the original "as a service" offering: Every major cloud service provider—Amazon Web Services, Huawei Cloud, Digital Ocean, Google Cloud, IBM Cloud, Microsoft Azure—started by offering some form of IaaS.

2.2.1 Components of IaaS

IaaS includes set of bodily and virtualized resources that offer clients with the primary constructing blocks needed to run programs and workloads in the cloud.

2.2.1.1 Physical Data Centers:

IaaS companies will manipulate big statistics centers, generally around the sector, that contain the physical machines had to electricity the diverse layers of abstraction and that are made to be had to end customers over the net. In most IaaS fashions, quit customers do now not interact without delay with the physical infrastructure, however it's far furnished to them as a carrier.[14]

2.2.1.2 Compute:

IaaS is usually understood as virtualized computing assets, so for the functions of this article we will outline IaaS compute as a virtual system. Providers control hypervisors, and quit customers can then programmatically provision virtual "instances" with the desired amount of compute and reminiscence (and every now and then garage). Most carriers provide each CPU and GPU for different sorts of responsibilities. Cloud computing also usually comes with helping services which includes auto-scaling and load balancing that provide the dimensions and performance characteristics that make the cloud suitable inside the first place.[15]

2.2.1.3 Networking

Cloud networking is a form of software-described networking wherein traditional community hardware, such as routers and switches, is made to be had programmatically, usually through an API. More advanced networking use cases include the construction of multi-zone areas and digital personal clouds, each of so that you can be discussed in more element later.[16]

2.2.1.4 Storage

The three number one kinds of cloud garage are “block garage", “record storage”, and item garage. Block and report storage are common in traditional information facilities, but can often war with the scale, normal overall performance, and distributed tendencies of the cloud. Object garage has as a result emerge as the maximum not unusual form of cloud storage of the three because it is appreciably allotted (and therefore durable), makes use of commodity hardware, facts may be without difficulty

accessed over HTTP, and scale isn't exceptional essentially unlimited, however performance scales linearly with growth cluster [17]

2.3 Virtualization

Processors, memory, storage, and other hardware components of a single computer can be separated into several virtual computer systems, or "virtual machines," using software to form an abstraction layer over the hardware (VMs). Even though each virtual machine (VM) only uses a portion of the actual underlying computer hardware, they each run their own operating system (OS) and operate like independent computers.[18]

The implication is that virtualization allows for greater green utilization of a computer's physical hardware and enables a greater return on business hardware investment. Nowadays, virtualization is a standard practice in enterprise IT architecture. It is also the methodology that drives the economics of cloud computing. Virtualization allows cloud companies to provide users with their current physical computer hardware; enables cloud customers to purchase the most convenient compute resources they need, when they want them, and scale those resources efficiently as their workloads evolve. [19]

Changing physical computers with software program-described VMs makes it less complicated to apply and manipulate regulations written in software program. This lets the user to create computerized IT provider management workflows. as an instance, computerized deployment and configuration gear allow directors to outline collections of digital machines and applications as offerings, in software templates. Which means they can installation the ones offerings time and again and constantly without cumbersome, time-eating. And error-prone guide setup. Admins can use virtualization security rules to mandate sure security configurations based at the function of the virtual machine. Regulations can even increase resource efficiency by way of retiring unused virtual machines to store on area and computing power.[20]

Replacing in-body computer systems with software-defined virtual machines makes it easier to apply and manage instructions written in software. This allows users to create computerized IT operator management workflows. For example, automatic deployment and configuration facilities allow directors to design collections of virtual machines and packages as services in software templates. This means they are able to set up their offers over and over and over without being bulky and time-consuming. And error-prone manual setup. Administrators can use virtualization security

guidelines to mandate certain security configurations based on the position of the virtual gadget. Regulations can even increase resource performance by stopping unused virtual machines from buying space and compute electricity.

2.4 Hypervisor

The hypervisor is the software layer that abstracts the hardware from the working machine permitting more than one operating structure to run at the identical hardware. The hypervisor runs at the host device permitting digital machines to run on the host's hardware as properly. [21]

A hypervisor, additionally known as a virtual device supervisor (VMM), manages these virtual machines as they run aspect through facet. It logically separates VMs from every other and assigns every its personal part of the underlying processing power, memory and garage. This prevents VMs from interfering with each other; so if, as an example, one OS suffers a crash or safety compromise, the others will survive. A hypervisor is a application that allows a couple of running systems to proportion a unmarried hardware host. Each operating machine appears to have the host processor, memory, and different assets all to itself. However, the hypervisor absolutely manages the host processor and resources, allocating what is wanted to every operating gadget in turn and making sure that the guest operating systems (referred to as virtual machines) can't intervene with each different.[22], [23]

2.4.1 Types of Hypervisors

The hypervisors can be broadly divided into two categories. Type 1 and Type 2.

2.4.1.1 Type 1 hypervisor

A Type 1 hypervisor runs directly on the host computer's physical hardware and interacts directly with its CPU, memory, and physical storage. For this reason, Type 1 hypervisors are also referred to as metal hypervisors. A Type 1 hypervisor replaces the host operating system. [24]

2.4.1.2 Type 2 Hypervisor

A Type 2 hypervisor does not run at once at the underlying hardware. Instead, it runs as an software inside the OS. Type 2 hypervisors are not often visible in server environments. Instead, they may be suitable for individual PC customers who want to run more than one working structures. Examples include engineers, security experts

studying malware, and enterprise users who need access to programs best available on different software program structures.[25]

2.5 Virtual Machine

“A virtual machine is a virtual representation or emulation of a physical computer.” They are frequently called a guest, at the same time as the bodily machine they run on is called a Physical Machine or guest. Virtualization lets in person to create more than one virtual machines, every with its personal working system (OS) and applications, on a unmarried physical machine. A digital pc can not talk at once with a bodily laptop. Instead, it wishes a light-weight software program layer called a hypervisor to coordinate between it and the underlying physical hardware. The hypervisor allocates physical computing sources—along with processors, memory, and storage—to every virtual system. It maintains each VM separate from the others so that they don't affect each other. Although this era can cross through many names, inclusive of digital server, virtual server example (VSI), and digital personal server (VPS)[26]

Every VM has its committed running device managemnt system that operates it. (OS), structures and programs. Virtual machines isolation, configuration and aid allocation is basically carried out through the tool hypervisor which manages the connection between the VMs and the bodily machines. The benefits of datacentre virtualization are basically better physical belongings usage, much less electricity consumption, decrease infrastructure general price of possession (TCO), higher device availability, dynamic useful resource control, a good deal much less supplier lock-in and adoption to cloud computing. These essential advantages to the IT enterprise purpose having extra than eighty percentage of the worldwide facts centers using digitalization technologies. [27]

2.5.1 Architecture of a Virtual Machine

A VM (virtual machine) is an emulation of a computer machine wherein those machines use pc architectures to offer the capability of a bodily computer. The physical tool on which the digital machines run is referred to as the Host, even as the digital machines are known as the Guest. One host may have a couple of guests.

Each digital hardware tool performs the equal characteristic for the digital gadget as the hardware at the physical laptop. Each virtual device has CPU, memory, and disk sources. CPU virtualization emphasizes overall performance and runs without delay at

the CPU every time feasible. Basic physical sources are used whenever viable. The virtualization layer executes instructions handiest as had to make the digital machines paintings as if they were jogging without delay on the bodily system. [28], [29]

All the cutting-edge working structures offer guide for digital reminiscence, which lets in software program to use extra memory than the laptop physically has. Similarly, the ESXi hypervisor affords help for digital device reminiscence over committing, where the quantity of visitor reminiscence configured for all digital machines may be extra than the quantity of physical host reminiscence.

The following are the components of a virtual Machine

a. Virtual Processor (vCPU)

b. Virtual RAM (vRAM)

c. Virtual Storage

d. Virtual Network

2.5.1.1 Virtual Processor (vCPU)

"A virtual processor is a representation of a physical processor core for a logical partition operating system that uses shared processors."

When consumer installation and run an running system on a server that is not partitioned, the working gadget calculates the range of operations it may carry out concurrently based at the wide variety of processors on the server. For instance, if person install an working gadget on a server that has 8 processors and each processor can carry out two operations concurrently, the working gadget can carry out sixteen operations right away. In the equal way, when consumer deploy and run an running device on a logical partition that uses committed processors, the working gadget calculates the quantity of operations it is able to carry out concurrently primarily based on the quantity of dedicated processors which are assigned to the logical partition. In either case, the working device can without problems calculate how many operations it is able to carry out immediately through counting the full wide variety of processors to be had to it.[30]

However, when user install and run a system software on a logical partition that uses shared processors, the running device can not calculate the whole number of operations from the fractional range of processing devices which can be assigned to the logical partition. The server firmware should consequently constitute the computing energy to be had to the running system because the complete wide variety

of processors. This allows the operating device to calculate the wide variety of concurrent operations it could carry out. A digital processor is a illustration of a bodily processor for the operating gadget of a logical partition that makes use of shared processors. [30], [31]

2.5.1.2 Virtual Memory

One of the principle benefits of using a virtual memory version is that it permits the working system to allocate greater pages of virtual reminiscence to applications than it may physically lower back up using a vital garage that uses digital memory over the dedicate. This is facilitated via the usage of techniques including paging unused virtual memory to disk and copying pages on write, as well as noting that not all pages of memory are usually wished or used by a application at anybody time. These techniques aren't distinctive to running systems and are regularly utilized by hypervisors to allocate extra RAM to digital machines (VMs) and LPARs than is physically to be had on the host Server Machines. [28], [32]

2.5.1.3 Virtual Storage

Virtual Disks is an illusion created through the structure in that the system appears to have extra memory than it truly has. Each person or program is given an cope with space, and each address space contains the equal range of storage addresses. In truth, best the ones components of the deal with area which might be wanted at any given time are loaded into the imperative storage. Z/OS stores inactive portions of cope with spaces in auxiliary storage. Z/OS manages cope with. [17]

2.5.1.4 Virtual Network

Virtual networks permit communique among more than one computer systems, digital machines (VMs), virtual servers, or other devices in distinctive office and facts center places. While physical networks join computer systems the usage of cabling and different hardware, virtual networks expand these abilities by way of the usage of control software to connect computers and servers over the Internet. It makes use of virtualized variations of conventional networking tools which include switches and community adapters, making an allowance for more green routing and less complicated community configuration adjustments.

A digital community permits devices in lots of locations to feature with the same capabilities as a conventional bodily network. This lets in data facilities to be spread

over specific bodily places and offers network administrators new and greater green options, which include the capacity to without problems alter the network as wanted without having to interchange off or purchase additional hardware; greater flexibility in delivering the network to precise desires and packages; and the potential to transport workloads throughout the community infrastructure without compromising carrier, security and availability. [28], [33]

2.6 Public Cloud Provider

A public cloud provider owns and operates the data centres where the workloads of their clients are hosted. To ensure quick access to apps and data, service providers take care of all hardware and infrastructure maintenance and offer broadband network connectivity. The underlying virtualization software is also managed by the cloud service provider. The "utility" model is a computational adaptation of the public cloud paradigm. One of the most significant changes in enterprise computing history is the introduction and use of public cloud services. In a public cloud, a third-party service provider produces computing resources, which can range from fully functional enterprise infrastructures and development to individual virtual machines (VMs) and ready-to-use software applications. a platforms available to consumers over the general internet are platforms. These resources may be freely available, or access may be purchased using a pay-per-use or subscription-based pricing model. Users frequently start comparing cloud service providers based on their website's specifications before comparing prices.

However, as mentioned above, the specs of any given VM only give you a very high level view of its potential performance when running your workload.

The number of vCPUs (virtual CPUs), the physical server chip technology, the amount of RAM (random access memory), and the comparison of storage type and volume size are all factors that can be easily explored. However, in practice, performance does not correlate very well with the website specifications advertised by the provider. For a more complete comparison, we need to dig deeper into the prices.

According to Gartner's Report 2021, the IaaS industry is still expanding, and cloud-native architecture is already the standard for contemporary workloads. According to Gartner, Inc., the global Infrastructure as a Service (IaaS) market rose by 41.4% in 2021 to reach $90.9 billion, up from $64.3 billion in 2020. In 2021, Amazon

continued to dominate the IaaS industry, with Microsoft, Alibaba, Google, and Huawei following.

The three biggest players in the IaaS market, together with the international Chinese technology corporation Alibaba, are Amazon (Web Services), Microsoft (Azure), and Google (Compute Engine).

[34]

In the IaaS market, the largest companies by revenue are Amazon (Web Services), Microsoft (Azure) and Google (Compute Engine), as well as the Chinese multinational technology company Alibaba.

2.6.1 AWS (Amazon Web Services)

AWS remains the largest player in the market with maximum market share and provides various services ranging from infrastructure technologies like compute, storage, and databases—to emerging technologies like machine learning and artificial intelligence, data lakes and analytics, and the Internet of Things(IoT)[35]

2.6.1.1 EC2

Amazon Elastic Compute Cloud (Amazon EC2) is IaaS component of the AWS that provides scalable computing capacity in the Amazon Web Services (AWS) Cloud. The Virtual Machines provided by EC2 range from 0.5 GB T2 single core processors nano machines to large servers of 3904 GB RAM and 128 virtual core processors.[35]

2.6.1.2 t2 Micro Machine

t2 micro is the free tier EC2 machine provided by the AWS. It is the most widely used public cloud virtual machine. It is equipped with 1 GB RAM

t2 instances are a low-cost, general-purpose instance type that provides a base level of CPU performance with the ability to exceed the base level if needed.

It is hosted with High frequency Intel Xeon processors which have frequency of 3.3 GHz Intel Xeon Scalable processor (Haswell E5-2676 v3 or Broadwell E5-2686 v4) [36]

2.6.2 Google Cloud Platform

As per the Gartner’s report on cloud computing, GCP is placed just after third just after AWS and MS Azure.[34]

Google Cloud consists of a set of physical assets, such as computers and hard drives, and virtual resources, such as virtual machines (VMs), contained in Google data centers around the world. It offers big data, analytics, machine learning (ML), load balancing and scaling. It also fully encrypts data transfers and communication between data centers

It provides multiple type of Resources ranging from V.Ms to container based services. [37]

2.6.2.1 GCP Compute Engine

Google Compute Engine (GCE) is the Infrastructure as a Service (IaaS) component of Google Cloud Platform, which is built on the global infrastructure that runs Google search, Gmail, YouTube and other services. Google Compute Engine allows users to run virtual machines (VMs) on demand.[37]

It provides Predefined machine types which are ready-made virtual machine configurations with a specific amount of vCPU and memory to run applications quickly.

It also provides customizable VMs. Using custom machine types, user can create virtual machines with the optimal amount of CPU and memory for the workload. This allows user to adapt the infrastructure to workload. If requirements change, user can use stop/start to move workload to a smaller or larger Custom Machine Type instance or to a predefined configuration.[15]

GCE provides multiple series of series of Machines that includes E2, N1, N2, N2D, T and many more.

2.6.3 MS Azure

Azure is the second largest player in field of IaaS Service Provider after AWS. The cloud computing service operated by Microsoft for managing applications through data centers managed by Microsoft. It provides Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) and supports many different programming languages, tools, and frameworks, including both Microsoft-specific and third-party software and systems. Currently it is proving 200+ services. [38]

2.6.3.1 Azure Virtual Machines

Azure Virtual Machines in IaaS component of the MS Azure, Which is hosted on Intel Processor including Xeon.

Enables user-created virtual machines or Microsoft-provided virtual machines to run in Microsoft Azure. These are charged per minute. To provision a VM, the user will need a VHD for the boot process (here we used a standard VHD), a VM size, an extension, and some configuration parameters depending on the selected VHD (administrator account, DNS name, etc.).

Hard disks used in virtual machines are persistent in memory as blob pages. The virtual machine needs to be recreated to preserve all content (for example, in case of hardware failure). Content follows the same failover rules as all other content in block storage (triple storage of all content, optional geographic replication to remote data centers). Microsoft Azure uses standard virtual hard disks, so migrate existing virtual machines from on-premises Hyper-V to the cloud to run them as virtual machines, download VHDs from Microsoft Azure and run VHDs locally, which can happen in the corresponding Hyper-V environment IN. Various versions of Windows Server and some Linux distributions are supported as guest operating systems for virtual machines. Multiple server applications such as SQL Server, SharePoint Server, BizTalk Server and Active Directory are also supported. Applications from Oracle's technology portfolio (WebLogic, Java, Oracle DB, etc.)[39]

In this paper we had used an x86 Virtual Machine with single vCPU and 1 GB RAM and 8 GB Storage.

2.6.4 IBM Cloud Platform

The IBM Cloud platform combines Platform as a Service (PaaS) with Infrastructure as a Service (IaaS) to provide an integrated environment. The platform scales and supports both small development teams and organizations as well as large enterprises. Globally deployed in data centers around the world.

IBM Cloud provides solutions that enable higher levels of compliance, security and management, with proven architectural patterns and rapid delivery methods for running mission-critical workloads. Available in data centers worldwide, in 19 countries with multi-zone regions in the Americas, Europe, Asia and Australia.

IBM Cloud offers the most open and secure public cloud for businesses with a next-generation hybrid cloud platform, advanced data and AI capabilities, and deep enterprise knowledge across 20 industries. Solutions are available depending on the needs for working in the public cloud, on-premise or a combination.

IBM's Infrastructure as a Service (IaaS) is a type of cloud computing that allows user to allocate the compute, network, storage and security resources on demand. IBM Cloud's approach to IaaS allows user to scale and shrink resources as needed across more than 60 data centers around the world.[40]

IBM Provides three main type of Virtual Machine Instances.

Bare Metal Servers Hourly or monthly single tenant servers that are dedicated, including server resources, with other customers.

Virtual Servers Scalable virtual servers that are purchased with cores and memory allocation.

IBM Cloud for VMware Solutions Quickly and seamlessly integrate or migrate on-premises VMware workloads using VMware's scalable, secure, high-performance infrastructure and industry-leading hybrid virtualization technology.

In our research work we had used a 1GB RAM single core x86 architecture machine belonging to the IBM Virtual server category.

2.6.5 Oracle Cloud

Oracle Cloud Infrastructure launched in October 2016 with a single realm and core services across compute, storage and networking. Since then, Oracle Cloud has expanded to more than 70 services available in 29 cloud regions worldwide, with plans to reach a total of 38 regions by the end of 2021. OCI offers relational, OLAP, JSON and NoSQL databases, containers, Kubernetes, serverless capabilities, Spark, streaming, Jupyter notebooks, VMware – a range of cloud services essential for almost any workload. In 2020 alone, Oracle Cloud Infrastructure launched nearly 400 new services, features, and enhancements.

Compute is a Linux/Windows machine built in the cloud with a shape that we can use to deploy different services depending on the application requirements. Compute machines have an attached OCPU, memory, storage, and a boot volume that defines the shape of the instance. We can also mount and unmount storage volumes.

Oracle Cloud Infrastructure Compute allows the user to provision and manage single-tenant servers (dedicated servers also known as bare-metal) or multi-tenant virtual

machines using the same set of APIs. You can spin really flexible, independently serviced, hour-long bare servers in minutes. Oracle offers GPU-based processors from Intel and AMD and Nvidia.[41]

1) **Bare Metal (BM) instance**: In which you get direct access to the underlying hardware. It provides a dedicated physical server for the highest performance and strong isolation. Used for heavy workloads.

2) **Virtual Machine (VM) Instance**: A VM instance runs on top of Bare Metal hardware. On top of the bare metal server is a hypervisor that virtualizes it in smaller virtual machines. Virtual machines are ideal for running applications that do not require the performance and resources (CPU, memory, network bandwidth, storage) of an entire physical machine.

3) **Dedicated VM Host:** This is a combination of Bare Metal and Virtual Machine. In this case, multiple VMs are running on Bare Metal and the entire server is dedicated to a single host. Autoscaling and instance pool are not yet supported in these cases. [42]

CHAPTER 3
BENCHMARKING AND SURVEY ON BENCHMARKING TOOLS

A benchmark is commonly a program designed to evaluate the overall performance of hardware or software program. Benchmarks provide a way to evaluate the behaviour and velocity of various answers at the same time as performing the identical undertaking (or set of tasks). Subsequent outcomes may be used to select the fine answer for the undertaking. For instance, if want a pc that plays annoying multi-threaded tasks, evaluating the effects of benchmarks that carry out similarly on more than one computer systems (i.e. Those that use many CPU cores) will assist to recognize the hardware details. User can pick the right one without any know-how. [20]

Another important component affecting computing performance is the number of CPU cores a V.M has get admission to it. This isn't always a factor for all programs; however many modern applications help multi-threading. In impact, because of this the software and/or working machine is able to distribute the computing duties among multiple cores. One fantastic tip to improve the pc's overall performance is to healthy the wide variety of threads (i.e. Cores) an application can assist to the number of cores the V.M has access to.[43]

3.1 Components of Cloud Benchmarking

The foremost component of a Benchmarking is SUT (System Understand Test) is to be tested cloud environment. This includes all hardware, underlying software and management systems used for the cloud service. The actual SUT component set differs depending on whether it is a white-box or black-box cloud.

3.1.1 Computing power

When we take a look at uncooked computing strength, we are speak me approximately CPU and RAM. Performance variations at a merely computational level between clouds are not clearly that huge. However, there are some elements that make actual variations.

By a ways of the most important element affecting computing performance within the cloud is competition. Public clouds are multi-tenant environments. RAM and garage cannot definitely be reallocated (although they can be resold), however CPU can and is. Contention tiers vary broadly, however in fashionable, public cloud providers are able to sell a bodily host's CPU capacity for extra than 100%.

3.1.2 Storage: the true measure of cloud server performance

All overall performance is restrained through the weakest hyperlink where in the bottleneck occurs. Currently, technology has advanced substantially within the discipline of virtualization with admire to CPU and RAM utilization. For example, one physical machine may be virtualized and have more than one cloud servers with minimal loss of basic combination performance. Sadly, there may be nevertheless a number of progresses to be made within the case of garage. The stop result is that, in most instances, the overall performance of virtual servers within the cloud is decided by way of the performance of the cloud storage Space.

3.1.3 Networking

Determining and measuring network performance is significantly easier than computing and disk efficiency. Network working - analysis has two key aspects, namely latency and bandwidth.

Depending at the needs, the latency of the network the service provider uses might also or won't depend. If use the cloud for mainly stand-by myself operations, latency is not likely to be a priority. However, if going for walks actual-time applications that speak with the world outdoor the cloud, latency may be a critical performance element.

3.2 Intel Benchmark Methodologies

The Intel MPI Benchmark runs a set of MPI performance measurements over various message sizes for point-to point and global communications.

Intel uses 3 sets of his MPI benchmarks. H. MPI-1, MPI-2, MPI-3. MPI-1 includes a set of tasks that perform single and parallel transfer operations . MPI-2 includes operations such as one-way communication, input/output benchmarking. MPI 3 includes Non-Blocking Collective (NBC) operations. These benchmarks takes account of the remote memory access (RMA) functionality introduced in the MPI-3

standard. IMB-MT - Benchmarking MPI-1 functions running in multiple threads per layer.

Sophisticated methods have been developed to analyze CPU performance using telemetry data from the performance monitoring units (PMUs) they capture basic micro architectural events using hardware counters. These events can be used to understand system behavior and application impact on various aspects of the CPU such as cache, memory and translation look aside buffers (TLBs). automation vendors (EDA). Intel also ensures that functionality is preserved and Communication patterns must guarantee a deterministic order of sending and receiving data. For point-to-point MPI-1 communication, user can do this by separating thread message flows with tags. However, this method cannot be used in his MPI-1 shared communication. So different methods are used for both collective and point-to-point benchmarks, each thread uses its own MPI communicator. Thread control within the ranks is performed using the OpenMP* API. [44]

3.3 AMD Benchmarking Methodologies

AMD generally uses two types of benchmarking methods: application-based benchmarks and synthetic benchmarks. App benchmarks show how optimized a platform is, but a particular version of a particular application isn't necessarily good at predicting how a new application will perform. Next, AMD also uses third-party tools such as his Cinebench, Geekbench, PC Mark, Passmark 10 and Blender Bench.There are two types of benchmarks which are commonly preffered for evaluation of Computer system performance: "Synthetic" and "Application-based". It can be both useful in the decision-making process, although individual benchmarks can often have undesirable properties. As a general rule the joint use of multiple benchmarks can alleviate these problems, provides a more reliable picture of performance.[45], [46]

with current performance. Even with benchmarks that use the installed version applications can have dramatic differences if they are not the same application versions are used on all tested systems. While app benchmarks show how well the platform is optimized they are not always good for specific versions of certain applications a predictor of what the new application's performance will look like. Unlike application benchmarks, synthetic benchmarks measure the overall performance potential of a particular platform. For example, many video conferencing

solutions use multiple CPU cores to perform functions such as virtual backgrounds. Synthetic Bench - markers that measure a platform's ability to use multiple threads and help predict how well a platform can provide this new feature.[46]

3.4 Spack Benchmarking Methods

Spack is an open-source tool that gives a bundle management framework and a device for installing complicated technical software program. It is designed to aid more than one software program versions and configurations across many distinctive structures and environments. Spack helps huge supercomputing facilities composed of many users and alertness teams. They percentage common software program installations on clusters with complex architectures using libraries with out a standard utility binary interface (ABI). Many configurations can coexist at the same device because installing a new version of Spack will no longer destroy current installations.

Spack offers a easy specification syntax that permits customers to exactly configure variations and alternatives. It simplifies the work of package deal authors because bundle files are written in Python™. The spec also lets in them to keep a single document for distinctive bundle builds.

AMD supports AMD Optimized CPU Compilers and Libraries (AOCC and AOCL) with Spack packages. AMD additionally supports Spack applications for typically used HPC benchmarks and a growing catalog of open-source medical programs with endorsed command line instructions the use of AOCC and AOCL. Each software-precise web page info the Spack command line, alternatives, and example commands for strolling the application.[47]

3.5 HPL Benchmark

HPL is a software package for solving (random) dense linear systems in double-precision (64-bit) arithmetic on distributed-memory computers. It can therefore be considered a portable and freely available implementation of the High-Performance Computing Linpack benchmark. The algorithms used in HPL can be summarized by the following keywords. Combining Column Broadcasts - Various Virtual Panel Broadcast Topologies - Swap Broadcast Algorithm to Reduce Bandwidth - Reverse Permutation with Depth 1 Lookahead. As the time required to calculate it. The best performance this software can achieve on the system depends on many factors. Nevertheless, the algorithms and accompanying implementations described here are

scalable due to some restrictive assumptions about the interconnection network in the sense that the parallel efficiency with respect to memory usage per processor remains [48]

3.6 HPCG Benchmarking Methods

The High Performance Conjugate Gradient (HPCG) Benchmark is designed as a successor to the HPL benchmark. The motivation behind developing the brand new benchmark become to create a memory performance metric for HPC systems that better displays the overall performance of not unusual application workloads on modern-day architectures. HPC workloads are now not limited handiest through CPU overall performance, but also can be touchy to reminiscence overall performance. Examples of reminiscence sensitivity are computational fluid dynamics programs inclusive of the open source OPENFOAM utility. While HPL solves a machine of linear equations, the HPCG benchmark focuses on data get admission to patterns..[49]

3.7 STREAM Benchmark

The STREAM benchmark is a simple artificial benchmark program that measures the sustained reminiscence bandwidth (MB/s) and corresponding computational velocity of a simple vector core. The general rule for strolling STREAM is that every array should be as a minimum 4 times the blended size of all last-stage caches used in the run, or 1 million elements, which ever is larger.

The STREAM benchmark is a simple synthetic benchmark software that measures the sustained memory bandwidth (MB/s) and corresponding computational pace of a easy vector core. [50]

3.8 SPEC Cloud IaaS 2018 benchmark

The 2018 SPEC Cloud® IaaS Benchmark measures the overall performance of infrastructure-as-a-provider (IaaS) cloud implementations. It helps trying out each public and private clouds.

The benchmark emphasizes the provisioning and runtime aspects of the cloud the use of I/O and CPU extensive cloud computing workloads. SPEC selected NoSQL database transaction on social networks and K-Means clustering the usage of map/reduce as two substantial and representative types of cloud computing workloads. The first workload makes use of the Yahoo Cloud Serving benchmark with

a NoSQL Cassandra database. The 2nd activity makes use of KMeans from the Intel Hibench Suite for Hadoop.

Each workload runs as a distributed application composed of 6 or 7 instances, referred to as application instances (AI). Benchmark provides multiple application instances at runtime. Adding AI to the cloud puts a strain on available cloud resources (eg CPU, memory, disk and network).[46], [47]

3.9 Cinebench Benchmark Method

Cinebench is a hardware evaluation benchmark especially popular in the field of 3D rendering. It is developed by Maxon, which also developed the 3D modeling suite Cinema 4D and GPU renderer Redshift. Major CPU manufacturers such as Intel and AMD use Cinebench scores as a benchmark and proof of quality for new high-end models. At Megaender, we use these scores as the basis for our calculator to help user to determine the cost of future projects.[51]

3.10 Geekbench Benchmarking Method

Geekbench measures CPU's single-core and multi-core performance for everything from checking email to taking a picture to playing music or everything at once. The Geekbench 5 CPU benchmark measures performance in new application areas including augmented reality and machine learning, so user will know how close the system is to peak.CPU Benchmark scores are used to evaluate and optimize CPU and memory performance using workloads that include data compression, image processing, machine learning, and physics simulation. Performance on these workloads is important for a wide range applications including web browsers, image editors, and developer tools.[52].[53]

Geekbench 5 provides two composite scores: single-core and multi-core. These scores are calculated using the weighted arithmetic mean of subsection scores. Subsection points are calculated using the geometric mean of the scores of the workloads contained therein subsection.

Geekbench generates one and overall scores for the Compute Benchmark. The total score is geometric mean score of individual computing loads. Each Compute task has an implementation for each supported Compute API. While it is possible to compare scores across APIs (eg OpenCL scores with Metal scores). Important to keep in mind that due to the nature of the Compute API there is a difference in performance may be

caused by more than differences in the underlying hardware (e.g. the GPU driver may have large impact on performance).

3.11 VMmark Benchmark

VMmark is a free tool used by hardware producers and others to examine the performance, scalability and electricity utilisation of virtualization platforms.

VMmark Benchmark Provides accurate and reliable benchmarking of digital information center performance and strength intake. Compare the overall performance and energy intake of different virtualization systems. Can be used to determine the overall performance impact of hardware, software program, or configuration changes inside a virtualized environment. Cloud environments typically aggregate numerous exclusive workloads onto a virtualized platform, a collection of physical servers with get admission to to shared garage and network resources. Traditional performance and scalability benchmarks for person workloads in non-virtualized environments had been not designed with virtual machines or cloud environments in mind. Even single-server virtualization benchmarks do no longer completely capture the complexity of today's virtualized statistics facilities. The VMmark benchmark combines commonly virtualized programs into predefined packages referred to as "tiles". The number of VMmark tiles that a virtualization platform can run, in addition to the cumulative performance of these tiles and numerous platform-level workloads, decide the VMmark 3 score.[54].

3.12 PCMark 10 Benchmarks

PCMark benchmarks measure the PC's performance using tests based on real-world applications and activities. In PCMark 10, these tests include everyday tasks such as web browsing, video chat, typing and spreadsheets, photo and video editing, and gaming.

Run the main benchmark and get a PCMark 10 score that can be used to compare PC to other systems or see the impact of Windows updates and hardware upgrades.

PCMark 10 updates many of the tests in PCMark 8 and adds new ones as well. Thanks to new and improved workloads, the main PCMark 10 benchmark takes less than half the time of the equivalent test in PCMark 8.

Each benchmark run produces a high-level score, a medium-level test group score, and a low-level stress score. It can also compare results side by side in the app.

PCMark 10 shares the same user interface style as 3DMark and VRMark. The familiar layout makes it easy to start benchmarking with PCMark 10.[55]

3.13 3D Mark Benchmarks

3DMark is a device for measuring the performance of PCs and mobile gadgets. It incorporates many extraordinary benchmarks, every designed for a particular elegance hardware from smartphones to laptops to high performance gaming Pcs. 3DMark works by way of strolling extensive photos and computing tests. The extra effective hardware, the smoother the assessments can be. Don't be amazed if the framerates are low. 3DMark benchmarks are very traumatic. [56]

3.14 Novabench Benchmarks

Novabench is a comprehensive computer benchmark software for testing CPU, RAM, GPU and hard drives. It provides detailed specifications about all of them. It evaluates both the graphics and processing capabilities of the computer and assigns it a score. A series of multi-core tests, each targeting a specific CPU instruction pipeline. These include general purpose instructions, FPUs, and other instruction set extensions.
The graphics tests are designed to run on most graphics cards and integrated graphics released since 2012 (or possibly older). If the graphics capabilities of the system do not meet the requirements of the test, the test will be skipped and a score of 0 will be assigned. If the user knows that graphics tests are not supported or the system does not have graphics capabilities, they can skip them by selecting Run all tests without GPU from the Tests menu.[57]

3.15 Microbenchmarks

Microbenchmarks address the smallest of all benchmarks and are a very simple and without problems defined metric that tracks and measures the performance of a small and unique piece of code. Unlike benchmarking, that is the manner of strolling a pc application to assess how well the runtime is appearing. A microbenchmark always concerns a totally small amount of code. Therefore, their implementation is notably fast. Because of this need to make certain that person are the use of them inside the proper vicinity. Implementing microbenchmarks when they're no longer essential is a waste of time. This is why it is critical to validate the usefulness of any micro benchmark want to feature to the challenge. By a long way, the simplicity of

microbenchmarking is its largest benefit – and its biggest disadvantage. This simplicity lets in to slim down the additives and paths concerned and simplify the process of finding the root purpose of performance issues. It additionally manner that the checks that microbenchmarking makes use of are incredibly repeatable with fewer variables affecting outcomes. At the identical time, microbenchmarks are unusable in lots of situations because they may be "zoomed in" and precise. For example, if consumer want to get a top-level view of application's universal performance, the usage of a microbenchmark won't cut it. Therefore, choosing the proper scale for benchmarks is important to powerful testing.[58], [59] [60]

Speed is defined as a given number of operations per unit time, such as the number of double-precision multiplications a CPU can perform per second. Perhaps we would like to know the ``maximum'' speed at which the CPU can execute floating-point instructions under ideal circumstances. We might be even more interested in how floating-point speed in the ``real world'' depends on (for example) the size and location of the memory references being worked with.

Latency is the amount of time a CPU (or other subsystem) has to wait for a aid or provider to come to be available after it is requested, and is in gadgets of the inverse of speed—as an example, milliseconds according to disk are seeking. However, latency isn't always the inverse of speed, as latency is regularly very extraordinary for an remoted request and a streamed collection of identical requests.

Bandwidth is a special case of rate. It measures the ``facts consistent with unit time'' brought among subsystems (for instance between memory and CPU). Information within the context of computers is commonly records or code prepared as a byte move, so an average unit of bandwidth is probably megabytes in keeping with second.

3.16 Performance Analysis

Detailed performance analysis is a complex process, but the goal is very simple: to identify the root cause of the problem. To do this, we proceed in an orderly step-by-step manner to the home domain of the fault containing the root cause. Since this chapter is specifically about virtualization and cloud computing, we will discuss how these two domains affect this process. But first, let's look at the root cause analysis flow at a high level.

Decide whether the problem affects the entire application or only a specific transaction or transaction type. If it's the latter, then the first step is to isolate the type

Detailed performance evaluation is a complicated manner, however the intention is quite simple: to discover the foundation purpose of the hassle. To do this, we proceed in an orderly step-by means of-step way to the house domain of the fault containing the basis motive. Since this chapter is especially about virtualization and cloud computing, we can speak how those domains affect this process. But first, allow's have a look at the foundation cause evaluation go with the flow at a high level.

Decide whether the hassle affects the complete application or handiest a selected transaction or transaction kind. If it's the latter, then the first step is to isolate the type of transaction this is inflicting the slowdown. This sets the context for in addition investigation.

Identify the complex layer or levels. The hassle may additionally lie between the 2 layers (ie due to community latency).

To in addition isolate the hassle, take a look at that the environment itself isn't always having a poor impact, which includes CPU exhaustion or reminiscence throttling. This would include the whole thing out of doors of the app itself; for example suspension for waste series.

Finally, isolate the problem to a selected issue, technique, or carrier call inside the application. From this point, we will determine whether or not the foundation purpose is algorithmic, CPU-centric, the end result of excessive VM suspensions, or resulting from outside bottlenecks consisting of I/O, network, or locks (synchronization). This last step is the first step to a ability answer.

CHAPTER 4
LITERATURE SURVEY

The past two decades have seen an increase in number of online services. The Internet has changed from a mere useful tool to a dominant part of life and culture. To support this phenomenal growth, the handful of computers that were once used to serve the entire online community is now evolved into clouds comprising hundreds of lakhs of servers in stadium-sized data centers. Servers run around the clock, processing millions of requests and providing access to petabytes of data to users around the world. Driven by the ever-increasing demand for more servers and larger data centers, academia, and industry are directing significant efforts to improve the state-of-the-art server hardware architecture and software design. However, these efforts often encounter a number of obstacles that arise due to the difficulty of measuring the performance of server systems. While traditional benchmarks used to measure computer performance are widely accepted and easy to use, benchmarking server systems requires significantly more expertise. Server software is complex to configure and operate and requires tuning of the server operating system and server software to achieve peak performance. While typical datasets for traditional benchmarks can be identified intuitively, datasets for server systems are diverse and include not only the content of the data, but also the frequency and patterns of access to that data. Additionally, while performance of traditional benchmarks is easily defined as the time required to complete a unit of work, server performance quantification. It is inherently more demanding because it must consider quality of service (request latency) and not just peak raw throughput.

4.1 Previous Benchmarking Work

Blesson Varghese used the 6 methodology which divided the test into the two categories i.e capture attributes and group attributes. Capture attributes included the tests to find the computational ability of the machines whereas the group attributes included the inter process communications, transfers and similar tasks.[6]

Cloud assessment by application: Virtual machines are then ranked according to their empirical performance (in this paper performance assessment is time-sensitive to complete execution). Values of individual criteria for in −μ the evaluation

performance is normalized using $\bar{v}_{i,j} = (V_{i,j} - \mu_j)/\sigma_j$ where μ_j is the mean value of $v_{i,j}$ on m VM and σ_j is standard deviation in i,j over m VM. Normalized values are used to evaluate Mpi Vms.[6]

Aaron Paradowsk while comparing the V.Ms hosted on Cloud stack and open stack took in account the factors including Processing speed, time required to process certain number of instructions and load on the Processor and Memory.[61]

With the performance checking out having diagnosed a few regions of upgrades for CloudStack it's far important that those are fed back for investigation and development. With CloudStack being a community pushed platform for maintenance and refinements the proposed plan to have those further investigated and likely implemented in new revisions could be to submit a precis of the findings to the CloudStack community.

The principle identified trends outlined in each show that despite the fact that CPU and reminiscence configurations can also growth as machines are deployed those have little to no impact on the deployment and deletion instances on both structures. Tough power size however is in which both platforms overall performance show a clear dating in phrases of the time taken to complete the assignment when it comes to the dimensions of the tough disk. Because the hard disk length increases, so does the time taken to complete the challenge.

OpenStack proven a extra relative correlation to the length of the increase leading to a relative growth in time taken additionally. CloudStack alternatively, although time taken expanded with the elevated length of hard disk it did not continually growth at a relative rate to the hard power length increase. Similarly to this the challenge has now unfolded in addition avenues of studies precise around ways for enhancing the CloudStack structures performance inside the regions identified and also in benchmarking performance of different cloud systems [61]

Shrutika Dhargave in her paper titled “Evaluation of different Hypervisors Performances using Different Benchmarks” benchmarked hypervirsors using the application benchmarking methods including Hadoop Benchmark, SIGAR framework and GPU Pass-through Performance, FTP and HTTP approach. This method gives broader outlook of benchmarking however the lite V.Ms were untouched.[22]

The paper titled "Experimenting with application-based benchmarks across cloud providers through multi-cloud execution and modeling" by Athanasia Evangelinou presents a better picture of cloud performance analysis from different service

providers. Compares cloud service providers based on performance against a set of applications. Applies the process is described in order to achieve an optimal compromise between the parameters. Although this work is more advanced in the field of combined metric investigation, it also does not appear to include a mechanism for triggering and performing measurements. Skymark is a framework designed for analyzing the performance of IaaS environments. The framework consists of 2 components - Grenchmark and C-Meter. Grenchmark is responsible for generating and dispatching the workload, while C-Meter consists of a workload scheduler and dispatches the job to a cloud manager that manages various IaaS cloud plug-in architectures. Skymark focuses on cloud services with low performance parameters such as CPU, memory, etc. and not on core types of applications.[62]

CloudCmp provides a methodology and its goal is very similar to our approach to estimating the performance and cost of legacy cloud applications. A strong customer cloud customer can use the results to compare different providers and decide whether they should migrate to the cloud and which cloud provider is the best fit for their application. CloudCmp identifies a set of performance metrics relevant to application performance and cost, develops a benchmark job for each metric, runs the jobs on different providers, and compares.[63]

Papaer titled "Cloud Service Benchmarking" by David Bermbach is a client-side prospectus of cloud benchmarking and performance comparison, but the method implemented by David is also an application-based benchmarking that targets high-end machines, entry-level machines were ignored as usual.[64]

4.2 Research gap

1. Numbers of papers have been published on several cloud bench marking schemes, methodologies and tools, however their always remains a question of performance of light weight virtual machines which are most widely used.
2. The Public cloud vendors including AWS, Azure, GCP, IBM and Oracle provide a wide range of Micro Virtual Machines, however apart from pricing, the performance of the machine also remains a very important factor in growth of the business. However there is a need to analyse the performance of machines in this segment.
3. Majority of the tools such as Geekbench, Cinebench, PC Mark often crash while testing the small machines. Majority of the above mentioned tools execute workloads such as rendering, ray tracing or some other A.I based algorithms, but in case of micro virtual machines, these set of complex commands cannot be executed, thus either the program stops or the machine crashes.
4. Majority of the benchmarking researchers used application-based bench marking. In this case, If the application used for bench marking gets a version update, the results of benchmarking may vary.
5. Majority of the tools are designed for Server side benchmarking, meaning there by the tools such as VM Mark, SPEC, HPL needs access of the entire machines and cloud providers do not provide it to the clients.
6. Majority of the modern start ups , use basic machines to deploy standalone application, Thus there is need to tap the potential the low end machines too as they can save cost of the company.
7. The performance of the Virtual Machine is highly dependent on the underlying hardware, and Hypervisor, Thus there is need to analyse the performance of these too.

CHAPTER 5
PROPOSED METHODOLOGY

To compare the light weight virtual Machines provided by the different vendors we have used the following configuration of each having the 1 GB RAM, 10 GB Storage and CPUs with the x86 Architecture running Ubuntu 20.04 LTS Server versions.
We performed the similar set of tasks to the Virtual machines and monitored the three main attributes.

1. Time taken to complete the Task
2. Average load on the CPU during the task
3. Average load on the memory during the task

Based on the previous Benchmarking done we divided the tasks in two categories

1. Data Processing Ability
2. Computational Attributes

Description of the Tests Performed is mentioned in table below.

Table 5.1 Description of the Tests

Group 1	Data Processing Ability	Data Transferring Ability	1.Read / Write Operations 2. Copy Operation
		Data Encoding/ Decoding Ability	1. Compression 2. Extraction
Group 2	Computational Ability	Instruction Handling Ability	1. Encryption 2. Decryption 3. Hashing
		Arithmetic Calculation Ability	1. Integer Handling 2.Random Number Generation 3. Sorting Test
		Float Calculation Ability	1. Floating Point Test

5.1 Group 1: Data Processing Ability

This group of processes includes the processes that measure the ability of a machine to handle the basic operations on data including read write operations, copy operation, compression and extraction operation on the data.

5.1.1 Data Transferring Ability

Here we used a 1 MB file, and created 1024 copies of it. It performed 1024 read and write operations continuously.
Secondly, we copied the same file i.e 1 GB Data, from one folder to another. These test were used to find the ability of the machine to handle the read, write and copy instruction.

5.1.1.1 Read Write Operation

We used an image of size 1 MB and created 1024 copies of it using python file read write operations. The objective of the test was to test the ability of the machine to handle a bulk of read write instructions. At the end this operation generated the 1 GB data that was used to complete the further tasks.

5.1.1.2 Copy Operation

We copied a folder of size 1 GB (generated from the previous test) from one folder to another. The copy operation puts load on CPU, Memory and Storage. The objective of the test was to test the ability of the machine to transfer data.
Performance of machines in Copy Operation

5.1.2 Data Encoding/ Decoding Ability

We used 1 GB Data and performed the compression and extraction operations over it. During the operation we measured the time taken by the machine, the load on the processor and load on the memory.

5.1.2.1 Compression Test

Data compression is the reduction of the number of bits needed to represent data. Data compression can save storage capacity, speed up file transfers, and reduce storage hardware and network bandwidth costs. Here in this test we used 1 GB data, (consist of files of 1 MB * 1024)

5.1.2.2 Extraction

Extraction is the process of retrieving the compressed the data. The objective of this test is to check the performance of the of the machine against the extraction algorithms.

5.2 Group 2: Computational Ability

This group is further divided into three set of tests which includes Instruction handling ability, Arithmetic calculation ability and floating-point handling ability.

5.2.1 Instruction Handling Ability

This set included three tests, namely hashing, encryption and decryption. In this Test we used a 1 GB iso image and performed the above operations over it.

5.2.1.1 Hashing Test

A given key or string of characters is transformed into another value through the process of hashing. This is typically represented by a key or shorter value that makes it simpler to find or use the original string.

A hash value, also known as a hash, is created by a hash function and is based on a mathematical hashing method. A good hash always employs a one-way hashing technique to prevent the conversion of the hash back to the original key.

In this test, the processor and memory will be under load, and the machines' performance will be measured against the hashing algorithms.

5.2.1.2 Encryption Test

Information is encoded using the encryption method. Through this procedure, the information's original representation is changed into ciphertext, a different format. Only parties with the proper authorization should be able to convert the ciphertext back to plaintext and access the original data. While encryption by itself won't stop jamming, it will hinder a possible interceptor from understanding the message.

In This test we used a used 1 GB iso image and encrypted it. This is used to measure the efficiency of the machine against the encryption algorithm. We used AES encryption for the purpose of this test.

5.2.1.3 Decryption Test

Decryption is the process of restoring encrypted data to its original state. This is typically a reverse encryption procedure. It decodes data that has been encrypted. We used the encrypted data produced in the previous step in this step.

5.2.2 Arithmetic Calculation Ability

This set of tests included the operations like Sorting, finding random integer values in a given range and calculating the prime numbers in the certain range. This Test puts the load on memory as well as Processor. This test ultimately test the calculation ability of a machine.

5.2.2.1 Sorting Test

Rearranging an array or list of elements according to an element comparison operator is done using a sort algorithm. The matching data structure's members are placed in a new order using the comparison operator. The Method used in this thesis is Merge Sort algorithm. Here we used a list of 1000000 integers and measured the impact of the test on the machine.

5.2.2.2 Random Number Generation Test

There are numerous uses for random number generators. The technique of frequently utilising a random number generator (RNG) to produce a series of numbers or symbols that cannot be predictably produced better than by chance is known as random number generation. This implies that a certain sequence that results will have some patterns.

5.2.2.3 Integer Test

In this test, we gave the machines algorithms to process which included the operations like addition, multiplication and division. The objective of the test was to find the arithmetic computational ability of the machines.

5.2.3 Float Calculation Ability

Floating-point (FP) arithmetic is arithmetic that represents real numbers approximately using a fixed-precision integer, called the significant, scaled by the exponent of a fixed-base integer. Floating point refers to the fact that the radix points

of a number can "float" anywhere to the left, right, or between the significant digits of the number.

In this test the machines were put under the test to perform operation over the floating points. This included operations of addition, multiplication and division of floating-point numbers.

CHAPTER 6
IMPLEMENTATION

For the performance analysis of the virtual machines, we chose the Machines belonging to the categories of Micro and small from different public cloud vendors. The Vendors included AWS, GCP, MS Azure, IBM Cloud and Oracle Cloud. To provide a similar running environment we chose the almost similar configuration of the machines hosted over different CPUs. However, the architecture of the Processor was chosen same i.e x86.We chose the machines with x86 architecture having 1 GB RAM and single virtual processing unit.

Following table 7.1 represents the configuration of the virtual machines under the Benchmarking test.

Table 6.2 Description of Machines Under Test

S.NO	Service Provider	Machine Series	Architecture	Hosted Over	No. of Virtual Cores	RAM (in GBs)	Cost per hour
1.	AWS	T2 Micro	x86	Xeon E5-2676 v3	1	1	$0.0052
2.	Azure	VM	x86	Xeon Platinum 8272CL	1	1	$0.0092
3.	GCP	N1	x86	Xeon(R) CPU	1	1	$0.03
4.	GCP	N2	x86	Xeon(R) CPU	2	1	$0.06
5.	GCP	E2 Micro	x86	Xeon(R) CPU	2	1	$0.0098
6.	GCP	N2D	x86	EPYC 7B12	2	1	$0.05
7.	IBM	Virtual Server C1.1X1	x86	Xeon(R) E5-2650 v4	1	1	$0.038
8.	Oracle	VM.Standard.E2.1.Micro	x86	EPYC 7551	1	1	0
9.	Oracle	VM.Standard3.Flex (Intel)	x86	Xeon(R) Platinum 8358	1	1	$0.0012
10.	Oracle	VM.Standard.E2.1.Micro (Prev Gen)	x86	AMD EPYC 7551	1	1	0

6.1 Implementation Algorithms

After Selecting the Virtual Machines with the above mentioned configuration the following tests were operated over each machines for five times.

Step 1

Run Setup Script

Step 2

Run Read Write Test Function

Step 3

Run Copy Test Function-Script

Step 4

Run Compression Test Function-Script

Step 5

Run Extraction Test Function-Script

Step 6

Download an ISO Image for further Test

Step 7

Run Hashing Test Function-Script on ISO image

Step 8

Run Encryption Test Function-Script on ISO image

Step 9

Run Decryption Test Function-Script on ISO image

Step 10

Run Sorting Test Function-Script on a list of 100000 integers

Step 11

Run Random Number Generation Test Function-Script

Step 12

Run Integer Test Function-Script

Step 13

Run Floating Test Function-Script

Step 14

Process the generated Data-Script

Step 15

Remove the files generated During Test

Step 16

Re Execute the steps 1 to steps 15 for 5 times

Step 17

Re organize the data generated in five rounds of tests

Step 18

Find Average of the five rounds of Tests

Step 19

Generate final Matrix

Step 20

Exit()

The following diagram figure 7.1 shows the flow of the work loads provided to the VMs in the given order.

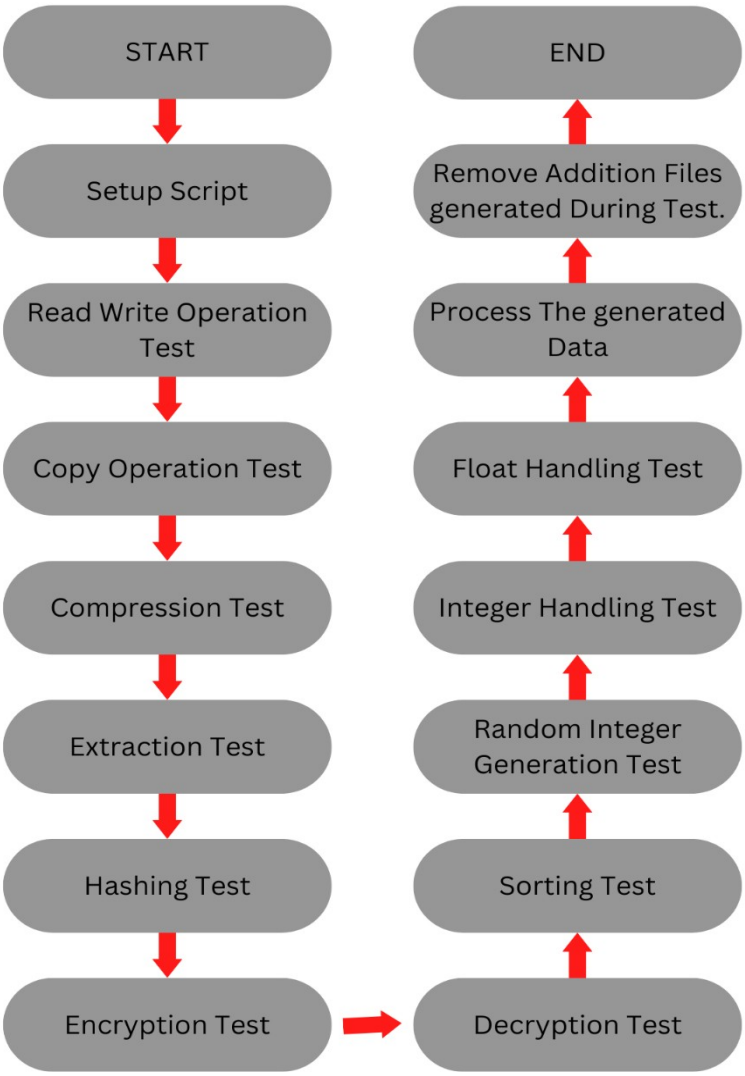

Figure 6.1 Order of Tasks Performed

CHAPTER 7

RESULTS AND ANALYSIS

7.1 Time based Comparison of Machines

Time remains the best factor to find efficiency of a machine, lower the time of execution, better the performance of the machine. The time of execution varied from process to process and machine to machine. Minimum time of execution for overall processes was by IBM C1.1x1 and GCP N2D machines.

Table 7.3 Time taken by machines in various Tests

MACHINE NAME	Data Transferring Ability	Data Encoding / Decoding Ability	Instruction Handling Ability	Arithmetic Calculation Ability	Float Calculation Ability
IBM C1.1X1	20.0131	30.6420	28.2748	93.2317	5.6761
GCP n1	28.4017	31.5857	29.7815	91.1323	5.9184
GCP -n2	28.1563	28.7914	28.7585	68.1720	3.9603
GCP -n2d	28.1563	28.7654	28.8080	63.4513	4.2616
GCP e2	38.9475	39.2135	35.4255	398.8322	5.3963
AWS T2 MICRO	63.0132	64.8042	63.5626	86.2745	6.3558
MS Azure	112.1442	151.802	124.1147	78.5818	N. A
oracle VM.E2.1.Micro	120.2799	161.309	129.7550	219.0820	12.1365
oracle E2.1.Micro (Prev Gen)	119.7574	160.892	130.2150	224.6971	11.5647
Oracle VM Standard3.Flex	119.9437	162.048	129.4551	53.2148	N. A

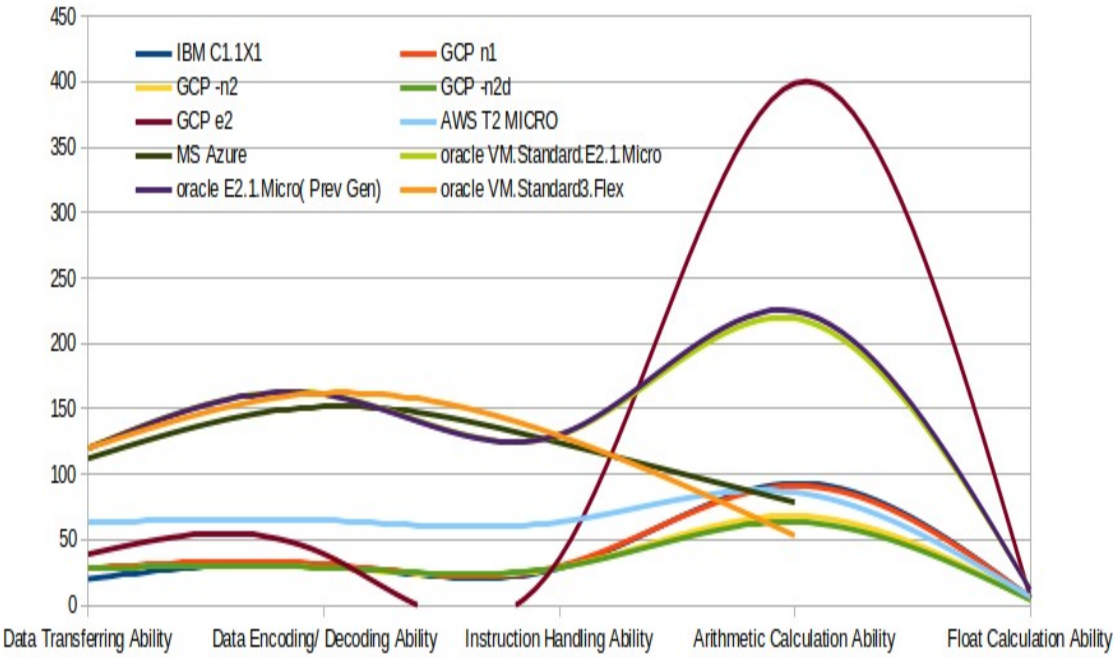

Figure 7.2 Time taken by Machines Per test

7.2 Load on the Processors of the machines during the entire Test

Each Process exerts different load on the processor, and each machine handles the load in the different manner, few machines put higher load on the processor to complete the task as early as possible and even reached to 99% of the efficiency to complete the test faster, few remained optimized and didn't put the load on the processor, thus took more time to complete the task.

MACHINE NAME	Data Transferring Ability	Data Encoding/ Decoding Ability	Instruction Handling Ability	Arithmetic Calculation Ability	Float Calculation Ability
IBM C1.1X1	2.686968	40.3598	40.35981	99.34728	89.96
GCP n1	2.444824	32.0838	32.08387	99.38707	82.95855
GCP -n2	0.700763	12.2590	12.25907	49.89512	45.6555
GCP -n2d	0.322567	12.3875	12.38759	49.93905	48.136
GCP e2	0.944037	13.0431	13.04318	16.40322	43.5924
AWS T2 MICRO	0.845794	16.2909	16.29095	99.60828	83.50225
MS Azure	0.658153	7.53203	7.532035	99.3006	N. A
oracle E2.Micro	0.086492	3.43551	3.435518	24.07081	21.73978
oracle E2.1.Micro (Prev Gen)	0.097034	3.36330	3.363304	23.7281	21.36302
oracle VM.Standard3.Flex	0.127231	2.53054	2.530542	49.79303	N.A

Table 7.4 Percentage of Processor utilized during the tests

The following figure represents the data in the graph format.

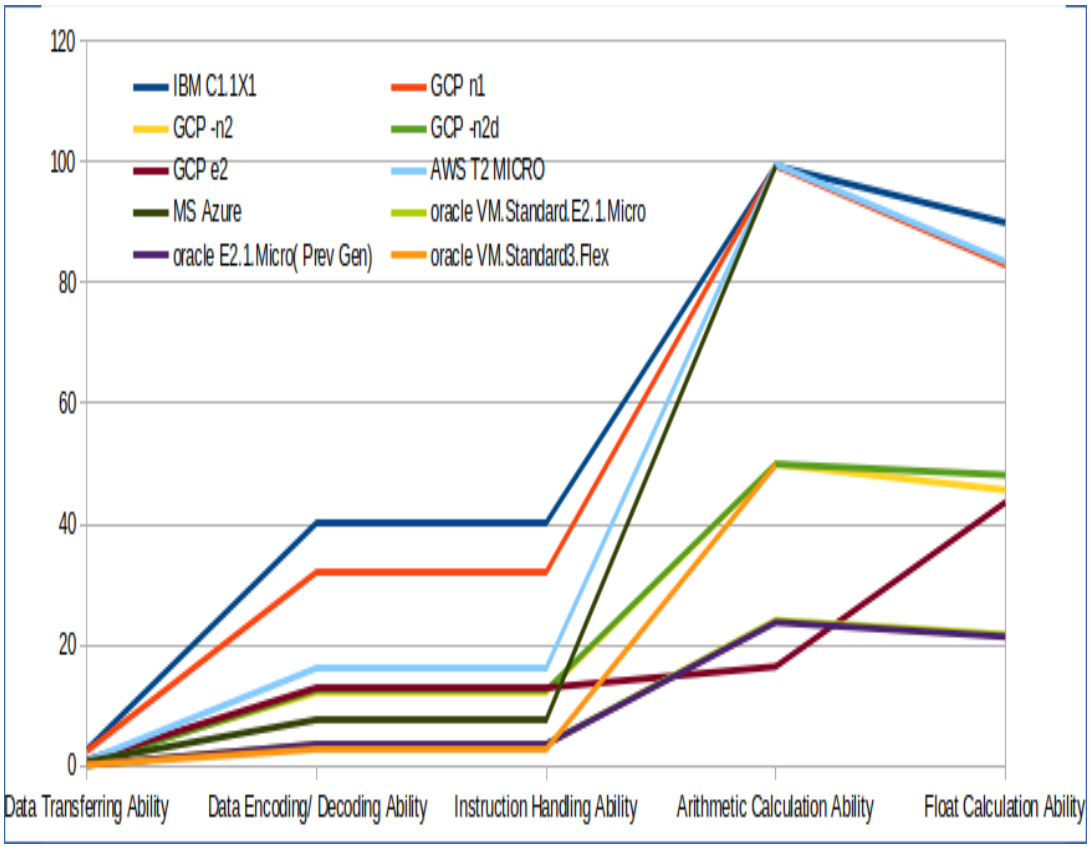

Figure 7.3 Load on the Processor During the Tests

7.3 Load on the Memory of the machines during the entire Test

Each process puts a different load on memory and each device handles the load in different manner. A few machines put a lot of load on the memory to do a task as fast as possible, and have even reached 60% efficiency to complete task faster.

Table 7.5 Percentage of Memory used during the tests

MACHINE NAME	Data Transferring Ability	Data Encoding/ Decoding Ability	Instruction Handling Ability	Arithmetic Calculation Ability	Float Calculation Ability
IBM C1.1X1	13.68465	13.2347	19.76469	24.24365	63.4336
GCP n1	12.93432	12.7573	19.19824	26.32538	66.760
GCP -n2	16.05344	16.0380	24.24912	27.90047	70.086
GCP -n2d	15.88931	15.8414	23.92122	27.75445	69.488
GCP e2	15.52807	15.4811	24.95081	26.28985	66.6036
AWS T2 MICRO	13.23983	13.0124	19.59634	26.07416	67.427
MS Azure	19.37303	19.2599	32.04859	34.29645	N.A
oracle E2.Micro	11.71429	11.5026	19.01387	24.30479	64.54
oracle E2.1.Micro (Prev Gen)	11.76423	11.5267	19.21813	24.22719	64.37
oracle VM.Standard3.Flex	20.48396	20.3075	33.83574	33.48337	N.A

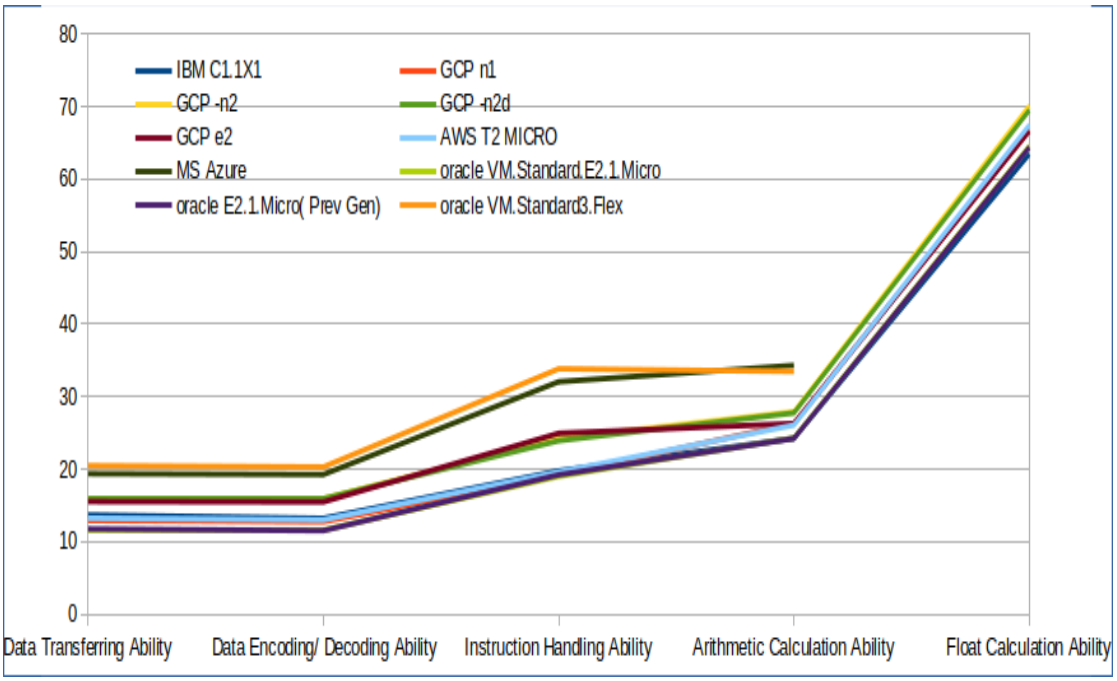

Figure 7.4 Load on the Memory During the Tests (in Percentage)

7.4 Final outcome

If a machine can perform a task faster than others, it should be called better and efficient, however even if a single task is not completed, the machines is said to be failed in the test. Thus, a machine which takes the lower time to complete the task should be called best. Irrespective of the load on the processor and memory, a user generally cares about the speed of the machines i.e how fast it completes the task. Thus, the Time should be considered as the best parameter to compare the final results. Two Machines namely machine by MS Azure and Oracle VM.Standard3.Flex were not able to complete the tasks. Thus, there time is not considered in outcome

Table 7.6 Total Time Taken to Complete the Test

S.No	MACHINE NAME	Total Time Taken
1.	**IBM C1.1X1**	177.8377
2.	**GCP n1**	186.8196
3.	**GCP -n2**	157.8385
4.	**GCP -n2d**	153.4426
5.	**GCP e2**	517.815
6.	**AWS T2 MICRO**	284.0103
7.	**oracle VM.Standard.E2.1.Micro**	642.5633
8.	**oracle E2.1.Micro(Prev Gen)**	647.1269

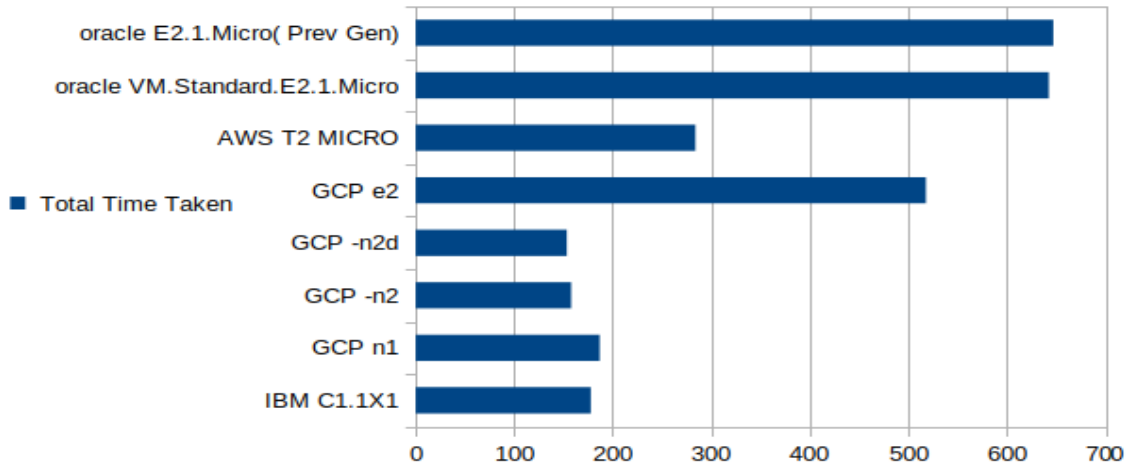

Figure 7.5 Total Time Taken To Complete the Test

CHAPTER 8
CONCLUSION AND FUTURE WORK

The performance of Machine be it real or virtual is dependent on multiple factors such as hardware, synchronisation between hardware and software. However factors like Processor and Memory play the deciding role. Secondly while hiring a cloud-based machine, factors such as availability, pricing per hour and Ease in deployment, operation and management cannot be ignored. This if take into account factors like price and ease, AWS proves to be best choice. However, oracle provides always free machines but performance is not up to the mark.

While majority of the benchmarking tools fail in testing the smaller machines and virtual machines can benchmark the machines in this segment. Moreover, it can also operate over a wide range of machines provided by different cloud vendors such AWS, MS Azure, Oracle cloud, GCP and IBM Cloud. While traditional benchmarking takes hours. Thus, it fills the gap in the category of benchmarking tools and provides an option for the benchmarking the lite V.Ms.

There is a need to compare the cloud platforms not only on the basis of number of services offered but also on the parameters such as pricing Vs Performance, performance of the different types of machines etc.

Light Machines in the categories of Micro and small having minimum configurations such as 1 GB RAM needs to be examined thoroughly.

Apart from the Traditional Cloud Platforms, the performance of emerging players can also be test, benchmarked and presented. Majority of the reports deals with the AWS, Azure and GCP, however performance and other features of Cloud Service Providers such as Alibaba, Oracle Cloud and IBM Public Cloud remains untouched. More over emerging players such as digital-ocean[65], linode[66], kamatera[67], Cloudsigma[68], Vultr Cloud[69] and many more can also be explored.

REFERENCES

[1] P. M. Mell and T. Grance, “The NIST definition of cloud computing,” Gaithersburg, MD, 2011. doi: 10.6028/NIST.SP.800-145.

[2] J. Scheuner and P. Leitner, “Performance benchmarking of infrastructure-as-a-service (IaaS) clouds with cloud workbench (Tutorial),” in *Proceedings - 2019 IEEE 4th International Workshops on Foundations and Applications of Self* Systems, FAS*W 2019*, Jun. 2019, pp. 257–258. doi: 10.1109/FAS-W.2019.00070.

[3] G. Phi, C. Tran, Y.-A. Chen, D.-I. Kang, J. P. Walters, and S. P. Crago, *Hypervisor Performance Analysis for Real-Time Workloads*.

[4] “Performance Analysis & Resolution of Cloud Applications | Dynatrace.” https://www.dynatrace.com/resources/ebooks/javabook/performance-analysis-and-resolution-and-a-cloud/ (accessed Oct. 11, 2022).

[5] “Benchmarking the cloud? New metrics required - Cloud computing news.” https://www.ibm.com/blogs/cloud-computing/2012/09/17/benchmarking-the-clouds-new-metrics-required/ (accessed Oct. 11, 2022).

[6] “Cloud computing | IBM.” https://www.ibm.com/cloud/learn/cloud-computing-gbl (accessed Oct. 10, 2022).

[7] “What is Public Cloud | IBM.” https://www.ibm.com/cloud/learn/public-cloud (accessed Aug. 26, 2022).

[8] N. K. Sehgal and P. C. P. Bhatt, “Features of Private and Public Clouds,” in *Cloud Computing*, Springer International Publishing, 2018, pp. 51–60. doi: 10.1007/978-3-319-77839-6_4.

[9] “What is Hybrid Cloud? | IBM.” https://www.ibm.com/cloud/learn/hybrid-cloud (accessed Aug. 26, 2022).

[10] “What are community clouds? - Cloud computing news.” https://www.ibm.com/blogs/cloud-computing/2014/08/27/community-clouds/ (accessed Oct. 10, 2022).

[11] “IAAS.” https://www.ibm.com/cloud/learn/iaas (accessed Aug. 26, 2022).

[12] “IaaS versus PaaS versus SaaS | IBM.” https://www.ibm.com/cloud/learn/iaas-paas-saas (accessed Oct. 10, 2022).

[13] “What is IaaS (Infrastructure-as-a-Service) | IBM.”

https://www.ibm.com/cloud/learn/iaas (accessed Sep. 28, 2022).

[14] “IaaS Cloud Computing Overview | Oracle.” https://www.oracle.com/cloud/what-is-cloud-computing/cloud-computing-overview/ (accessed Aug. 26, 2022).

[15] “What is Compute Engine? Use cases, security, pricing and more | Google Cloud Blog.” https://cloud.google.com/blog/topics/developers-practitioners/what-compute-engine-use-cases-security-pricing-and-more (accessed Sep. 29, 2022).

[16] “What is Virtual Networking? | VMware Glossary.” https://www.vmware.com/topics/glossary/content/virtual-networking.html (accessed Oct. 13, 2022).

[17] “How virtual storage addressing works in z/OS - IBM Documentation.” https://www.ibm.com/docs/en/zos-basic-skills?topic=storage-how-virtual-addressing-works-in-zos (accessed Oct. 13, 2022).

[18] I. Odun-Ayo, O. Ajayi, and C. Okereke, “Virtualization in cloud computing: Developments and trends,” in *Proceedings - 2017 International Conference on Next Generation Computing and Information Systems, ICNGCIS 2017*, Nov. 2018, pp. 35–37. doi: 10.1109/ICNGCIS.2017.10.

[19] “What is Virtualization? | IBM.” https://www.ibm.com/cloud/learn/virtualization-a-complete-guide (accessed Aug. 26, 2022).

[20] J. Fabo, “The impact of virtualization on benchmark accuracy,” MASARYK UNIVERSITY, 2020.

[21] “What is a hypervisor?” https://www.redhat.com/en/topics/virtualization/what-is-a-hypervisor (accessed Oct. 11, 2022).

[22] S. Dhargave, “Evaluation of different Hypervisors Performances using Different Benchmarks,” 2016. [Online]. Available: http://www.ijser.org

[23] B. DrAnupam Bhatia GurjeetsinghBhattal Research Scholar DCSA DCSA CRSU Jind and H. CRSU Jind, “A comparative study of Various Hypervisors Performance,” *Int. J. Sci. Eng. Res.*, vol. 7, 2016, [Online]. Available: http://www.ijser.org

[24] “Hypervisors | IBM.” https://www.ibm.com/cloud/learn/hypervisors (accessed Oct. 11, 2022).

[25] “What is a Bare Metal Hypervisor? | VMware Glossary.”

https://www.vmware.com/topics/glossary/content/bare-metal-hypervisor.html?resource=cat-1418329861#cat-1418329861 (accessed Oct. 14, 2022).

[26] “Virtual Machines | IBM.” https://www.ibm.com/cloud/learn/virtual-machines (accessed Oct. 10, 2022).

[27] M. E. Elsaid, H. M. Abbas, and C. Meinel, “Virtual machines pre-copy live migration cost modeling and prediction: a survey,” *Distrib. Parallel Databases*, pp. 1–34, Dec. 2021, doi: 10.1007/S10619-021-07387-2/FIGURES/7.

[28] J. E. Smith and R. Nair, “The architecture of virtual machines,” *Computer (Long. Beach. Calif).*, vol. 38, no. 5, pp. 32–38, 2005, doi: 10.1109/MC.2005.173.

[29] “Virtual Machine Components.” https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.vm_admin.doc/GUID-A70B7BF0-923E-443A-A1DD-11D99B6330EE.html (accessed Oct. 12, 2022).

[30] “Virtual processors - IBM Documentation.” https://www.ibm.com/docs/en/power8/8408-44E?topic=processors-virtual (accessed Oct. 12, 2022).

[31] “Demystifying the Number of vCPUs for Optimal Workload Performance,” 2016, Accessed: Oct. 12, 2022. [Online]. Available: https://d1.awsstatic.com/whitepapers/Demystifying_vCPUs.pdf

[32] “Virtual Memory Objects - IBM Documentation.” https://www.ibm.com/docs/en/aix/7.2?topic=interfaces-virtual-memory-objects (accessed Oct. 12, 2022).

[33] G. Somani and S. Chaudhary, “Application performance isolation in virtualization,” *IEEE Int. Conf. Cloud Comput.*, vol. 2009, pp. 41–48, 2009, doi: 10.1109/cloud.2009.78.

[34] “Gartner Says Worldwide IaaS Public Cloud Services Market Grew 41.4% in 2021.” https://www.gartner.com/en/newsroom/press-releases/2022-06-02-gartner-says-worldwide-iaas-public-cloud-services-market-grew-41-percent-in-2021 (accessed Aug. 26, 2022).

[35] A. Whitepaper, “Overview of Amazon Web Services.”

[36] “Amazon EC2 Instance Types - Amazon Web Services.” https://aws.amazon.com/ec2/instance-types/ (accessed Sep. 28, 2022).

[37] “Google Cloud overview | Overview.” https://cloud.google.com/docs/overview/ (accessed Aug. 26, 2022).

[38] “Azure services | Microsoft Docs.” https://docs.microsoft.com/en-us/azure/availability-zones/region-types-service-categories-azure (accessed Aug. 26, 2022).

[39] H. Sirtl, “Overview of Microsoft Azure,” 2015.

[40] “What is the IBM Cloud platform? | IBM Cloud Docs.” https://cloud.ibm.com/docs/overview?topic=overview-whatis-platform (accessed Nov. 01, 2022).

[41] Y. V Ravi Kumar, N. Basha, K. Kumar K M, B. M. Sharma, and K. Kerekovski, “Oracle Cloud Overview,” in *Oracle High Availability, Disaster Recovery, and Cloud Services*, Apress, 2019, pp. 517–574. doi: 10.1007/978-1-4842-4351-0_9.

[42] “Compute Shapes.” https://docs.oracle.com/en-us/iaas/Content/Compute/References/computeshapes.htm (accessed Nov. 02, 2022).

[43] “Benchmarking cloud servers: A Cloud Computing Insider’s Guide.” https://www.cloudsigma.com/benchmarking-cloud-servers-a-cloud-computing-insiders-guide/ (accessed Oct. 11, 2022).

[44] “Multithreaded MPI-1 Benchmarks.” https://www.intel.com/content/www/us/en/develop/documentation/imb-user-guide/top/multithreaded-mpi-1-benchmarks.html (accessed Oct. 15, 2022).

[45] “AMD CPU Benchmarks.” https://cpu-benchmarks.com/benchmarks/amd-cpu-benchmarks/ (accessed Sep. 29, 2022).

[46] “AMD Ryzen - Cpu Benchmark Specs & Test.” https://cpu-benchmark.org/amd-cpus/amd-ryzen/ (accessed Sep. 29, 2022).

[47] “Introduction to Spack - AMD.” https://developer.amd.com/spack/introduction/ (accessed Oct. 17, 2022).

[48] “HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed-Memory Computers.” https://netlib.org/benchmark/hpl/ (accessed Oct. 18, 2022).

[49] “HPCG - AMD.” https://developer.amd.com/spack/hpcg-benchmark/ (accessed Oct. 17, 2022).

[50] “STREAM Benchmark - AMD.” https://developer.amd.com/spack/stream-benchmark/ (accessed Oct. 18, 2022).

[51] “Cinebench Release 23 Installers.” https://www.maxon.net/en/downloads/cinebench-r23-downloads (accessed Sep.

30, 2022).

[52] Primate Labs, “Geekbench 5 CPU Workloads,” p. 14, [Online]. Available: https://www.geekbench.com/doc/geekbench4-cpu-workloads.pdf%0Ahttp://support.primatelabs.com/kb/geekbench/interpreting-geekbench-4-scores%0Ahttps://www.geekbench.com/

[53] “Geekbench 5 - Cross-Platform Benchmark.” https://www.geekbench.com/index.html (accessed Sep. 06, 2022).

[54] “VMmark Virtualization Benchmark.” https://www.vmware.com/products/vmmark.html (accessed Oct. 18, 2022).

[55] AMD, “AMD RYZEN ™ PRO 5000 SERIES PROCESSORS : MEASURING WHAT MATTERS – BENCHMARK CONSIDERATIONS FOR COMMERCIAL PC PURCHASES PERFORMANCE IS EVERYTHING.” pp. 1–12.

[56] “UL Benchmarks 3DMark user guide.” https://support.benchmarks.ul.com/support/solutions/articles/44002146295-3dmark-user-guide (accessed Oct. 18, 2022).

[57] “Novabench - Documentation - Test Information.” https://novabench.com/docs/tests#cpu (accessed Oct. 18, 2022).

[58] J. Scheuner and P. Leitner, “Estimating Cloud Application Performance Based on Micro-Benchmark Profiling,” *IEEE Int. Conf. Cloud Comput. CLOUD*, vol. 2018-July, pp. 90–97, 2018, doi: 10.1109/CLOUD.2018.00019.

[59] B. Varghese, L. T. Subba, L. Thai, and A. Barker, “DocLite: A Docker-Based Lightweight Cloud Benchmarking Tool,” in *Proceedings - 2016 16th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, CCGrid 2016*, Jul. 2016, pp. 213–222. doi: 10.1109/CCGrid.2016.14.

[60] “What is Micro-benchmarking?” https://www.adservio.fr/post/what-is-microbenchmarking (accessed Oct. 18, 2022).

[61] IEEE Computer Society *et al.*, “Benchmarking the performance of openstack and cloudstack,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7759 LNCS, no. February, pp. 173–188, 2015, doi: 10.1007/978-3-319-25414-2_14.

[62] A. Evangelinou *et al.*, “Experimenting with application-based benchmarks on different cloud providers via a multi-cloud execution and modeling framework,” *Commun. Comput. Inf. Sci.*, vol. 512, pp. 213–227, 2015, doi:

10.1007/978-3-319-25414-2_14.

[63] A. Li, X. Yang, S. Kandula, and M. Zhang, “CloudCmp: Comparing Public Cloud Providers,” 2010.

[64] D. Bermbach, E. Wittern, and S. Tai, *Cloud service benchmarking: Measuring quality of cloud services from a client perspective*. 2017. doi: 10.1007/978-3-319-55483-9.

[65] “DigitalOcean – The developer cloud.” https://www.digitalocean.com/ (accessed Sep. 06, 2022).

[66] “Cloud Computing & Linux Servers | Alternative to AWS | Linode.” https://www.linode.com/ (accessed Sep. 06, 2022).

[67] “Kamatera – Performance Cloud Infrastructure.” https://www.kamatera.com/ (accessed Sep. 06, 2022).

[68] “Cloud servers & Hosting | Switzerland Europe US APAC.” https://www.cloudsigma.com/ (accessed Nov. 11, 2022).

[69] “SSD VPS Servers, Cloud Servers and Cloud Hosting - Vultr.com.” https://www.vultr.com/ (accessed Nov. 11, 2022).

LIST OF PUBLICATION

1. G. Yugansh and G. Bhavana, “A Comprehensive Survey Of Infrastructure As A Service By Top Public Cloud Vendors,” *Int. J. Creat. Res. THOUGHTS - IJCRT*, vol. 10, no. 11, pp. 888–895, 2022, doi: http://doi.one/10.1729/Journal.32205.
2. G. Yugansh and G. Bhavana “Performance Analysis and Comparisonof the Micro Vrtual Machines Provided by the Cloud Vendors “ Paper presented in the “World Conference on Science Engineering and Technology(WCSET) held in Bhopal, India on 02 -03 December, 2022”

   The presented paper will be published in Scopus indexed journals

PLAGIARISM REPORT

Ouriginal by Turnitin

Document Information

Analyzed document	thesis-7-1-23-IEEE-FORMAT.docx (D155672212)
Submitted	1/13/2023 5:32:00 AM
Submitted by	Bhavana
Submitter email	bhavanagupta@sistec.ac.in
Similarity	2%
Analysis address	bhavanagupta.rgpv@analysis.urkund.com

Sources included in the report

W	URL: https://bisite.usal.es/archivos/cloud_computing_and_multiagent_systems_0.pdf Fetched: 1/19/2022 6:05:32 AM	1
W	URL: https://mu.ac.in/wp-content/uploads/2021/01/Cloud-Computing.pdf Fetched: 10/9/2021 1:59:34 PM	2
W	URL: https://topic.alibabacloud.com/article/3-models-of-cloud-computing-services-iaas-paas-saas_1_1... Fetched: 2/1/2021 10:33:19 AM	5
W	URL: https://www.researchgate.net/publication/317719529_A_Review-Cloud_and_Cloud_Security Fetched: 2/17/2020 2:30:42 PM	1
W	URL: https://ijcsit.com/docs/Volume%207/vol7issue2/ijcsit2016070231.pdf Fetched: 4/4/2022 8:58:35 AM	1

Entire Document