PROJECT REPORT

"DESIGN A LOAD BALANCING MODULE IN PRIVATE CLOUD USING A META-HEURISTIC ALGORITHM"

By

GYANESH GUNJAN	10300117088
DHIRAJ YADAV	10300117093
ABHISHEK ANAND	10300317123

UNDER THE GUIDANCE OF

Dr. Tarun Kumar Ghosh
(Professor, Haldia Institute of Technology)

REPORT SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE AND ENGINEERING

HALDIA INSTITUTE OF TECHNOLOGY

Haldia, Purba Medinipur, West Bengal

Affiliated to Maulana Abul Kalam Azad University Of Technology

Approved by AICTE

HALDIA INSTITUTE OF TECHNOLOGY

(Department of Computer Science and Engineering)

Project Title

"DESIGN A LOAD BALANCING MODULE IN PRIVATE CLOUD USING A META-HEURISTIC ALGORITHM"

By

GYANESH GUNJAN *(17/CSE/088)*

DHIRAJ KUMAR YADAV *(17/CSE/093)*

ABHISHEK ANAND *(B18/CSE/140)*

ACKNOWLEDGEMENT

We would like to express our special thanks of gratitude to our mentor ***Dr. Tarun Kumar Ghosh***, who gave us the excellent opportunity to do this wonderful project on the topic ***Design a load balancing module in private Cloud Using a meta-heuristic algorithm***. They helped us in our Research and we came to know about so many new things. The blessing, help and guidance given by him from time to time will carry me a long way in the journey of life on which we are about to embark.

We are obliged to our project team members for the valuable information provided by them in their respective fields. We are grateful for their cooperation during the period of assignment.

PROJECT TEAM INFORMATION

Name : Gyanesh Gunjan
Address :
WARD NO-08, VILL/POST-CHATAUNI,PO-BASOPATTI, DIST-MADHUBANI, Bihar (847225)
Email : gyaneshgunjan777@gmail.com
Contact No. : 9534554600

- -

Name : Dhiraj Kumar Yadav
Address :
MQ 46, BANDH COLONY, KATHARA, BOKARO, JHARKHAND (PIN:829116)
Email : dhiraj.yadav2873@gmail.com
Contact No. : 8340799159

- -

Name : Abhishek Anand
Address :
OLD EXCHANGE ROAD WARD NO 10, HOUSE NO 65, SITAMARHI,BIHAR(843302)
Email : abhishekrockstoread@gmail.com
Contact No. : 6204733017

- : CONTENTS : -

ABSTRACT

The extensive use of computers has raised a large number of real-life optimization problems in different fields that are complex and difficult to solve. Metaheuristics define algorithmic frameworks that can be applied to solve such problems in an approximate way, by combining constructive methods with local and population-based search strategies, as well as strategies for escaping local optima.
Cloud computing strives to provide computational resources according to user's demand through the internet on pay per use. The required resources may be software, physical or hardware infrastructure. The chief focus of cloud computing is load balancing-the process of assigning load over multiple virtual servers to achieve user satisfaction as well as optimal resource utilization. Several static and dynamic algorithms have been formulated that efficiently manage the load balancing in a cloud environment.
In this paper, a load balancing algorithm based on honey bee behavior (LBA_HB) is proposed. Its main goal is to distribute workload of multiple network links in a way that avoids underutilization and over utilization of the resources.
This can be achieved by allocating the incoming task to a virtual machine (VM) which meets two conditions:
1. number of tasks currently processing by this VM is less than number of tasks currently processing by other VMs
2. Available processing power of the current system is better than other systems in the network.
The proposed algorithm is compared with single system execution and response time. The results of experiments show the efficiency of the proposed algorithm in terms of execution time, response time.

Keywords :
Cloud Computing, Virtualization, Load Balancing, Honey Bee Foraging Behaviour, Throughput

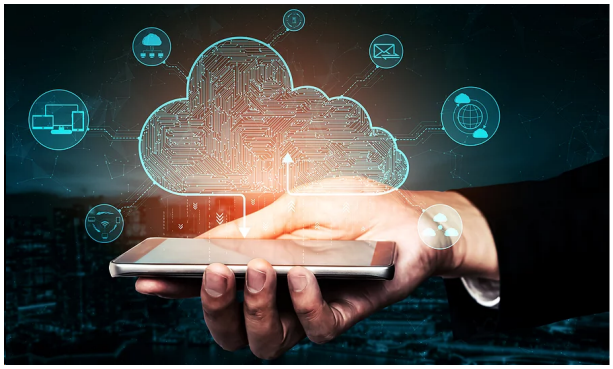
Introduction

Cloud computing, often referred to as simply “the cloud,” is the delivery of on-demand computing resources — everything from applications to data centers — over the internet on a pay-for-use basis, without direct active management by the users.
Cloud computing is one of the most emerging technologies that provides standards for large scale distributed and parallel computing. It’s a framework for enabling applications to run on virtualized resources and accessed by common network protocol and standards.It provides computing and infrastructural resources and services in a very flexible manner that can be scaled up or down according to demand of the end user. The cloud can reduce the information system’s cost through virtualization technologies. Cloud computing is a new computer model whose purpose is to suggest the IT services into requested services (on-demand). The goal is to allow services to be accessible from anywhere, anytime, and by anyone (country as you go).

Cloud load balancing is the process of distributing workloads across multiple computing resources. Cloud load balancing reduces costs associated with document management systems and maximizes availability of resources. Load balancing is the process of redistribution of workload in a distributed system like cloud computing ensuring no computing machine is overloaded, under-loaded or idle. Load balancing tries to speed up different constrained parameters like response time, execution time, system stability etc. thereby improving performance of cloud. It is an optimization technique in which task scheduling is an NP hard problem. There are a large number of load balancing approaches proposed by researchers where most of focus has been concerned on task scheduling, task allocation, resource scheduling, resource allocation, and resource management.

Cloud Computing

According to the National Institute of Standards and Technologies (NIST), cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Being a cloud computing provider doesn’t mean just supplementing IT resources, it means providing strategic, core information technology.




(figure source - Internet)

Characteristics of cloud computing

There are basically 5 essential characteristics of cloud computing.

1. **On-Demand self-services** :

The Cloud computing services does not require any human administrators, users themselves are able to provision, monitor and manage computing resources as needed. It is one of the significant and essential features of Cloud Computing. It enables the client to constantly monitor the server uptime, abilities, and allotted network storage. This is a fundamental characteristic of Cloud Computing, and a client can likewise control the computing abilities as per his needs.

2. **Cloud Broad Network Access**

We'll have a look at the National Institute of Standards and Technologies (NIST) definition first :

“Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms e.g. mobile phones, tablets, laptops and workstations”.

~ NIST

The cloud characteristic of Broad Network Access means that the service can be accessed from multiple locations (such as from a corporate office or from home) using multiple different types of client (such as a Windows PC or an Android mobile).

3. Rapid elasticity

A key characteristic and benefit of cloud computing is its rapid scalability. This cloud characteristic enables cost-effective running of workloads that require a vast number of servers but only for a short period. Many clients have such workloads, which can be run very cost-effectively because of the rapid scalability of Cloud Computing. The Computing services should have IT resources that are able to scale out and in quickly and on as needed basis. Whenever the user requires services it is provided to him and it scales out as soon as its requirement gets over.

4. Resource pooling

The IT resources (e.g., networks, servers, storage, applications, and services) present are shared across multiple applications and occupant in an uncommitted manner. Multiple clients are provided service from the same physical resource. Resource pooling is one of the essential characteristics of Cloud Computing. Resource pooling means that a cloud service provider can share resources among several clients, providing everyone with a different set of services as per their requirements. It is a multi-client strategy that can be applied to data storage services, processing services, and bandwidth provided services. The administration process of allocating resources in real-time doesn't conflict with the client's experience.

5. Measured services

National Institute of Standards and Technologies (NIST) states that

"Cloud systems automatically control and optimize resource use by leveraging and metering capability at some level of abstraction appropriate to the type of service. For example, storage, processing, bandwidth, and active user accounts.
Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service."

The resource utilization is tracked for each application and occupant, it will provide both the user and the resource provider with an account of what has been used. This is done for various reasons like monitoring billing and effective use of resources. Cloud services can be accessed by users in a Pay-per-Use-On-Demand model in which guaranteed high quality services are offered to the customer by the service providers. Consumers could be charged fees which may be in terms of hours, data transfers or other use-based attributes delivered.

Cloud Service Models

The term "cloud service" refers to a wide range of services delivered on demand to companies and customers over the internet. These services are designed to provide easy, affordable access to applications and resources, without the need for internal infrastructure or hardware. From checking email to collaborating on documents, most employees use cloud services throughout the workday, whether they're aware of it or not.

There are the following three types of cloud service models.

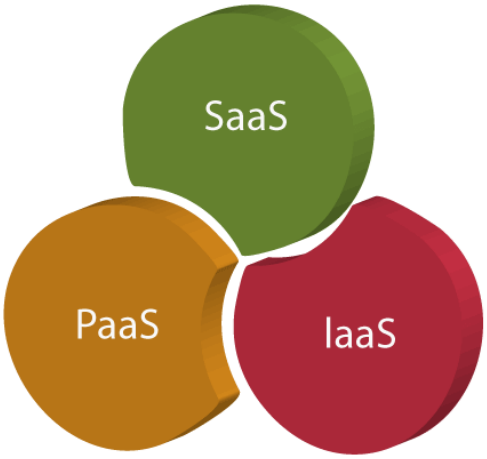


Figure source url

1. Software as a Service (SaaS)

Software as a service (or SaaS) is a way of delivering applications over the Internet—as a service. Instead of installing and maintaining software, you simply access it via the Internet, freeing yourself from complex software and hardware management. It removes the need to install and run applications on our own computers or in the data centers eliminating the expenses of hardware as well as software maintenance.
Software as a service (or SaaS) provides a complete software solution which you purchase on a **pay-as-you-go** basis from a cloud service provider.Most SaaS applications can be run directly from a web browser without any downloads or installations required.The SaaS applications are sometimes called **Web-based software**, **on-demand software**, or **hosted software**.

Characteristics of SaaS

- **Cost Effective -** *Pay only for what you use*
- **Reduced Time -** *Users can run most SaaS apps directly from their web browser without needing to download and install any software.*
- **Accessibility -** *Accessible over the internet.*
- **Automatic Updates -** *Users are not responsible for hardware and software updates. Updates are applied automatically.*
- **Scalability -** *It allows the users to access the services and features on demand.*

Example: *BigCommerce, Google Apps, Salesforce, Dropbox, ZenDesk, Cisco WebEx, ZenDesk, Slack, and GoToMeeting.*

2. Infrastructure as a Service (IaaS)

Infrastructure as a service, or IaaS, provides the infrastructure that many cloud service providers need to manage SaaS tools—but don't want to maintain themselves. It serves as the complete data center framework, eliminating the need for resource-intensive, on-site installations. It is also known as **Hardware as a Service (HaaS).** The main advantage of using IaaS is that it helps users to avoid the cost and complexity of purchasing and managing the physical servers.

Characteristics of IaaS

- **Cost Effective -** *Eliminates capital expense and reduces ongoing cost and IaaS customers pay on a per use basis, typically by the hour, week or month.*
- **Website hosting -** *Running websites using IaaS can be less expensive than traditional web hosting.*
- **Security -** *The IaaS Cloud Provider may provide better security than your existing software.*
- **Maintenance -** *There is no need to manage the underlying data center or the introduction of new releases of the development or underlying software.*

Example: *DigitalOcean, Linode, Amazon Web Services* (AWS), *Microsoft Azure, Google Compute Engine* (GCE), *Rackspace, and Cisco Metacloud.*

3. Platform as a Service (PaaS)

Platform as a service (PaaS) is a complete development and deployment environment in the cloud, with resources that enable you to deliver everything from simple cloud-based apps to sophisticated, cloud-enabled enterprise applications. You purchase the resources you need from a cloud service provider on a pay-as-you-go basis and access them over a secure Internet connection. A PaaS provider hosts the hardware and software on its own infrastructure. As a result, PaaS frees users from having to install in-house hardware and software to develop or run a new application.Thus, the development and deployment of the application takes place independent of the hardware.
The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

Characteristics of IaaS

- **Simple and convenient for users -** *It provides much of the infrastructure and other IT services, which users can access anywhere via a web browser.*
- **Cost Effective**
- **Efficiently managing the lifecycle -** *It is designed to support the complete web application lifecycle: building, testing, deploying, managing and updating.*
- **Efficiency -** *Support multiple languages and frameworks.*

Example: AWS *Elastic Beanstalk*, *Windows Azure*, *Heroku*, *Force.com*, *Google App Engine*, *Apache Stratos*, *Magento Commerce Cloud*, *and OpenShift.*

Cloud Deployment Models

National Institute of Standards and Technologies (NIST) defines four cloud deployment models: public clouds, private clouds, community clouds, and hybrid clouds. A cloud deployment model is defined according to where the infrastructure for the deployment resides and who has control over that infrastructure. Deciding which deployment model you will go with is one of the most important cloud deployment decisions you will make.

1. Public Cloud

Public Cloud is a type of cloud hosting that easily allows the accessibility of systems & its services to its clients/users. As the name suggests, this type of cloud deployment model supports all users who want to make use of a computing resource, such as hardware (OS, CPU, memory, storage) or software (application server, database) on a subscription basis. Most common uses of public clouds are for application development and testing, non-mission-critical tasks such as file-sharing, and e-mail service.
Some examples of companies that provide public cloud facilities are *IBM, Google, Amazon, Microsoft, etc.*
This type of cloud computing is a true specimen of cloud hosting, where the service providers render services to various clients. From the technical point of view, there is the least difference between private clouds and public clouds along with the structural design. Only the security level depends based on the service providers and the type of cloud clients use. The public cloud is better suited for business purposes for managing the load. This type of cloud is economical due to the decrease in capital overheads.

Characteristics of Public Cloud

- **Hassle-free infrastructure management**
- **High scalability**
- **Reliable**
- **Place independence**
- **Flexible**

The Disadvantages of a Public Cloud

- **Compromised reliability**
- **Data security and privacy issues give rise to concern**
- **The lack of a bespoke service**

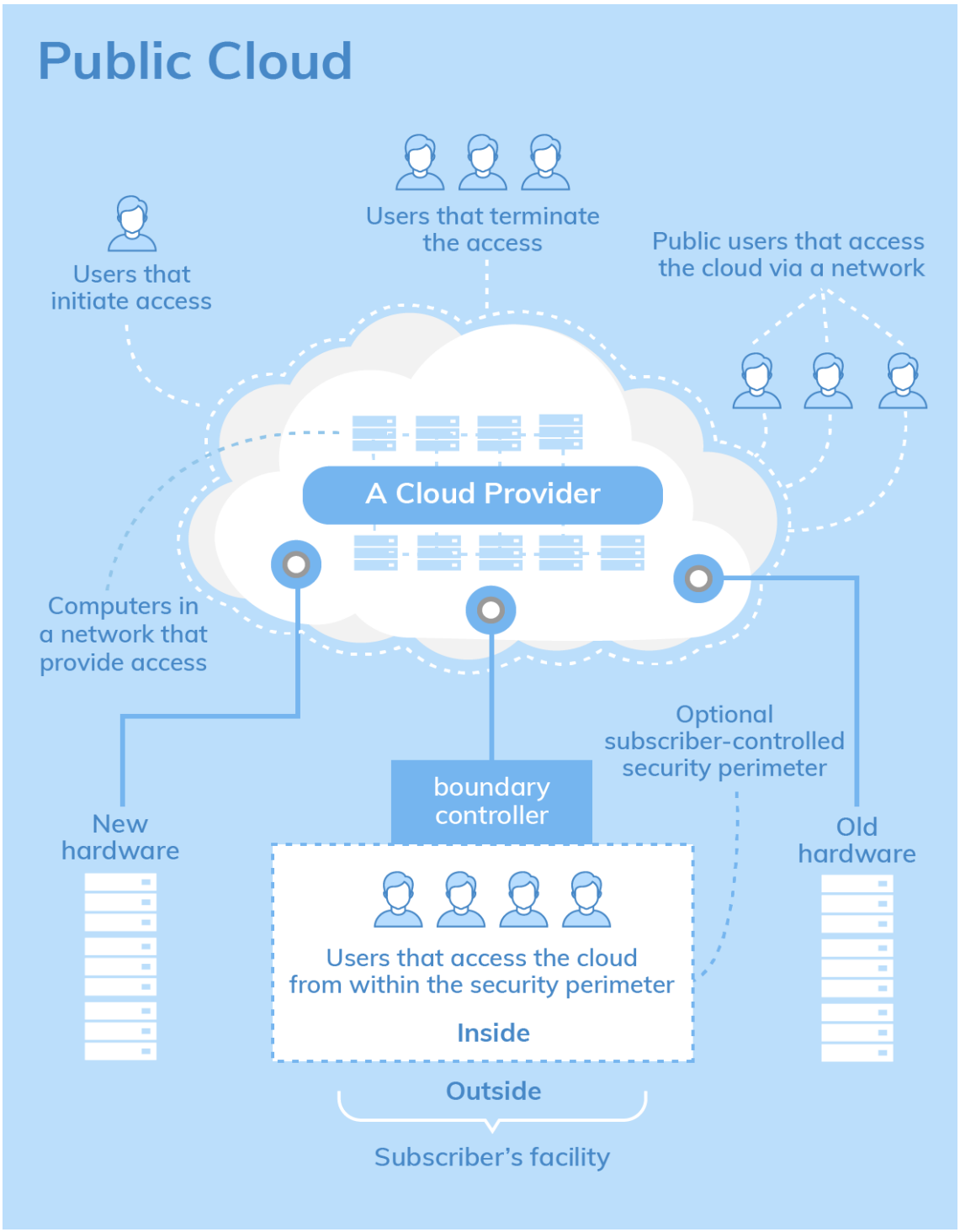


Figure source [url]

2. Private Cloud

Private Cloud is also termed as 'Internal Cloud', which allows the accessibility of systems and services within a specific boundary or organization. True to its name, a private cloud is typically infrastructure used by a single organization. Such infrastructure may be managed by the organization itself to support various user groups, or it could be managed by a service provider that takes care of it either on-site or off-site. Private clouds are more expensive than public clouds due to the capital expenditure involved in acquiring and maintaining them. Private clouds permit only authorized users, providing the organizations greater control over data and its security. However, private clouds are better able to address the security and privacy concerns of organizations today.

Characteristics of Private Cloud

- **Highly private and secured**
- **Control Oriented**

The Disadvantages of a Public Cloud

- **Poor scalability**
- **Costly**
- **Restriction**

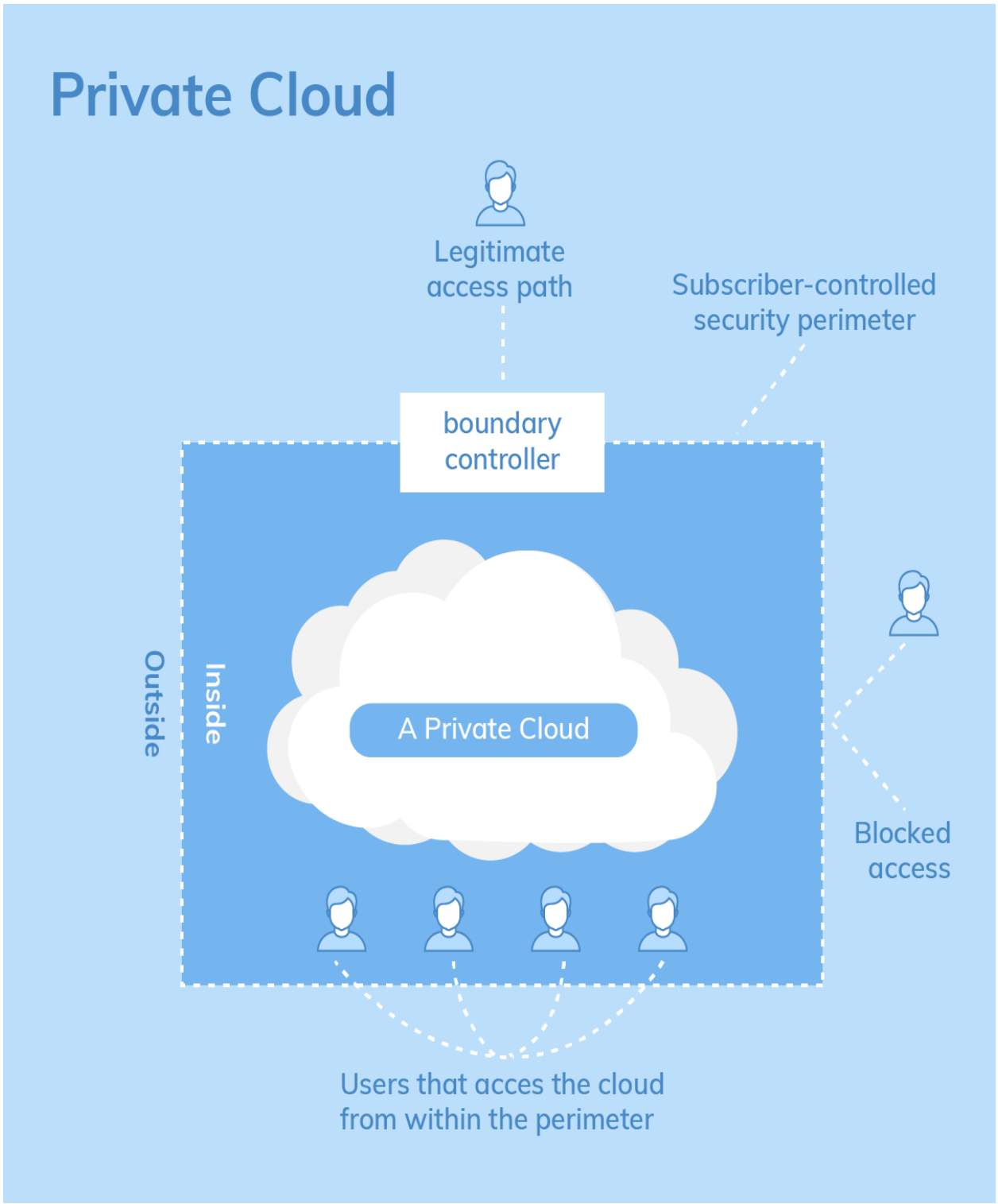


Figure source [url]

3. Community Cloud

Community Cloud is another type of cloud computing in which the cloud setup is shared manually among different organizations that belong to the same community or area. This deployment model supports multiple organizations sharing computing resources that are part of a community; examples include universities cooperating in certain areas of research, or police departments within a county or state sharing computing resources. Access to a community cloud environment is typically restricted to the members of the community.

Characteristics of Community Cloud

- **Cost reduction**
- **Improved security, privacy and reliability**
- **Ease of data sharing and collaboration**

The Disadvantages of a Community Cloud

- **High cost compared to the public deployment model**
- **Sharing of fixed storage and bandwidth capacity**
- **Not commonly used yet**

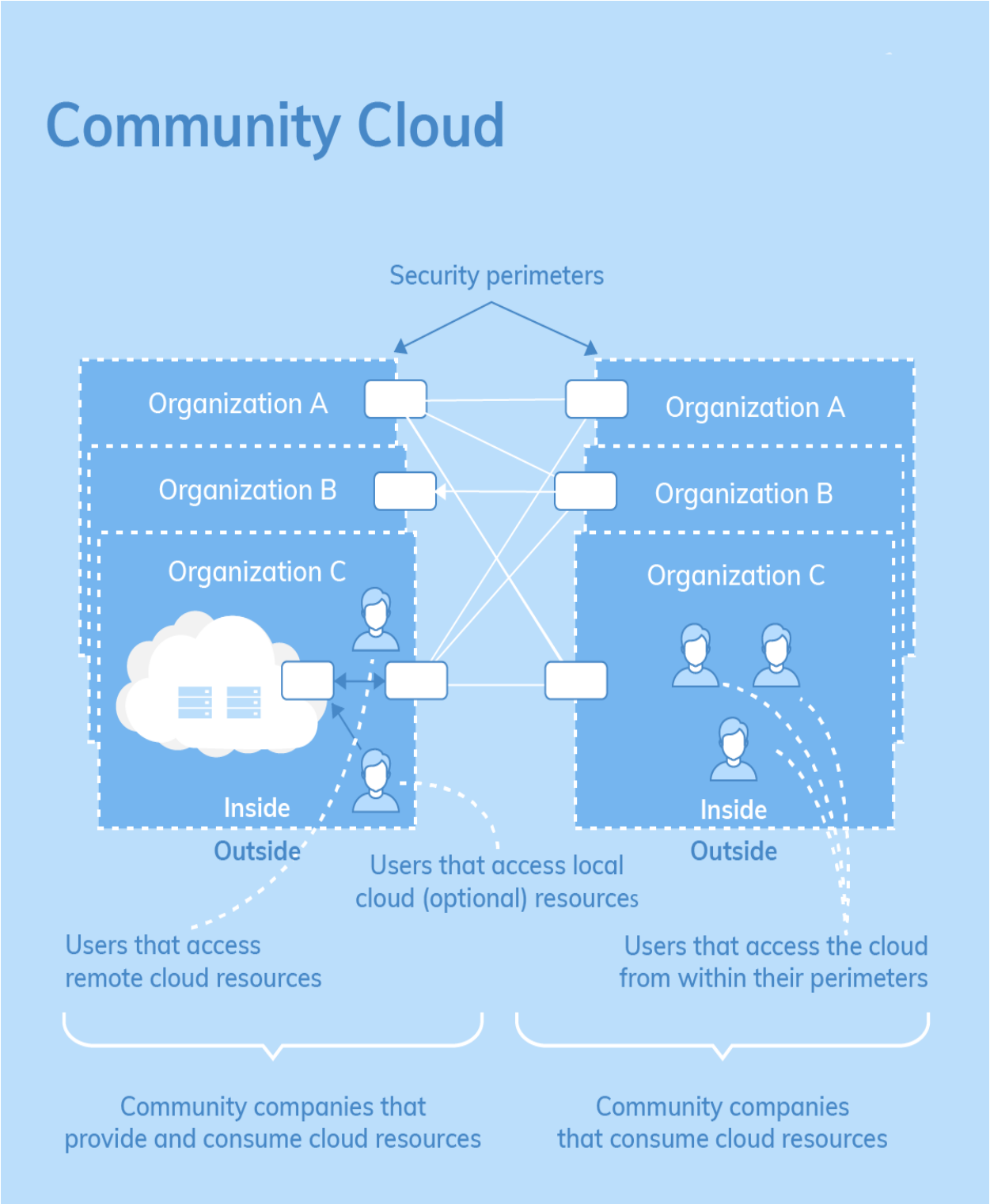


Figure source [url]

4. Hybrid Cloud

Hybrid Cloud is another cloud computing type, which is integrated, i.e., it can be a combination of two or more cloud servers, i.e., private, public, or community combined as one architecture, but remain individual entities. Non-critical tasks such as development and test workloads can be done using the public cloud. In a hybrid cloud, an organization makes use of interconnected private and public cloud infrastructure. Many organizations make use of this model when they need to scale up their IT infrastructure rapidly, such as when leveraging public clouds to supplement the capacity available within a private cloud. For example, if an online retailer needs more computing resources to run its Web applications during the holiday season it may attain those resources via public clouds.

Characteristics of Hybrid Cloud

- **Improved security and privacy**
- **Enhanced scalability and flexibility**
- **Reasonable price**

The Disadvantages of a Hybrid Cloud

- **Complex networking problem**
- **Organization's security Compliance**

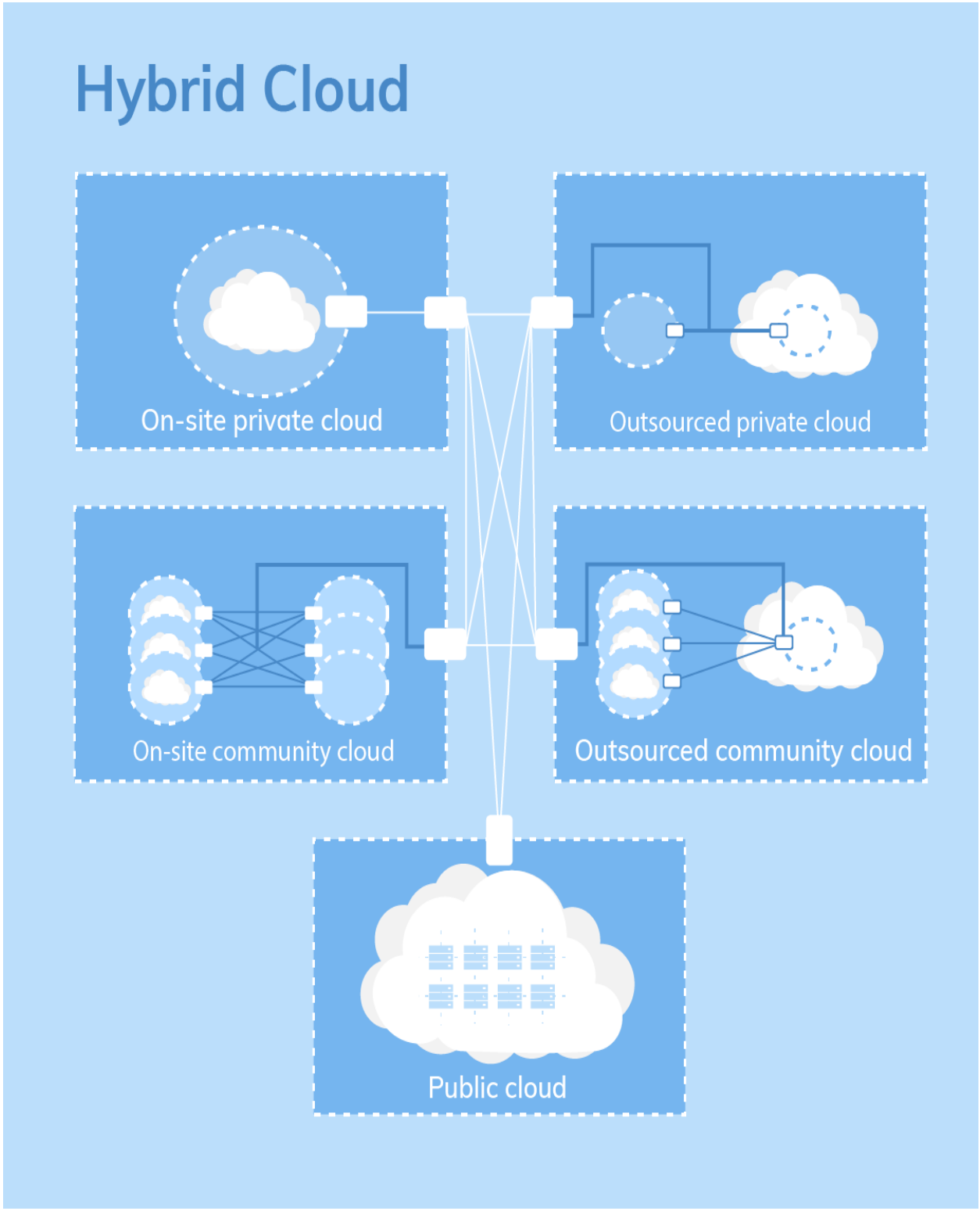


Figure source [url]

Virtualization

The term 'Virtualization' can be used in many respects for computers. It is the process of creating a virtual environment of something which may include hardware platforms, storage devices, OS, network resources, etc. It has become a critically important focus of the IT world in recent years. Virtualization technologies are used by countless thousands of companies to consolidate their workloads and to make their IT environments scalable and more flexible.
Virtualization is an essential characteristic of cloud computing, through which applications can be executed independently without regard for any particular configuration.Virtualization in cloud computing allows a provider to virtualize servers, storage, or other physical hardware or data center resources, which can then, in turn, allow them to provide numerous services such as infrastructure, software, and platforms. With cloud virtualization, customers can use only the parts of the service or resources they need, without needing the accompanying physical infrastructure themselves.

For example, *cloud virtualization technology for storage infrastructure allows users to access huge amounts of storage as they need it, and scale this need up and down as necessary, without needing any of their own storage.*

Virtualization authorizes multiple virtual machines with guest OS to run simultaneously & independently, on the same hardware machine as shown in Fig.

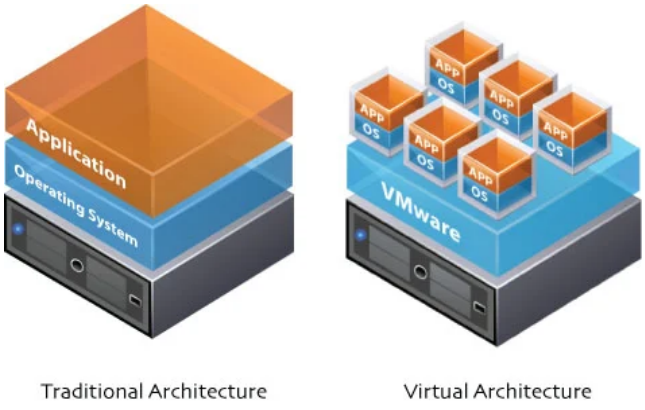


Figure Source Url

What Are The Benefits Of Virtualization?

Numerous benefits are provided by virtualization which includes, reduction in costs, efficient utilization of resources, better accessibility and minimization of risk among others.

Benefits for Companies

1. Removal of special hardware and utility requirements
2. Effective management of resources
3. Increased employee productivity as a result of better accessibility
4. Reduced risk of data loss, as data is backed up across multiple storage locations

Benefits for Data Centers

1. Maximization of server capabilities, thereby reducing maintenance and operation costs
2. Smaller footprint as a result of lower hardware, energy and manpower requirements

Types of Virtualization in Cloud Computing:

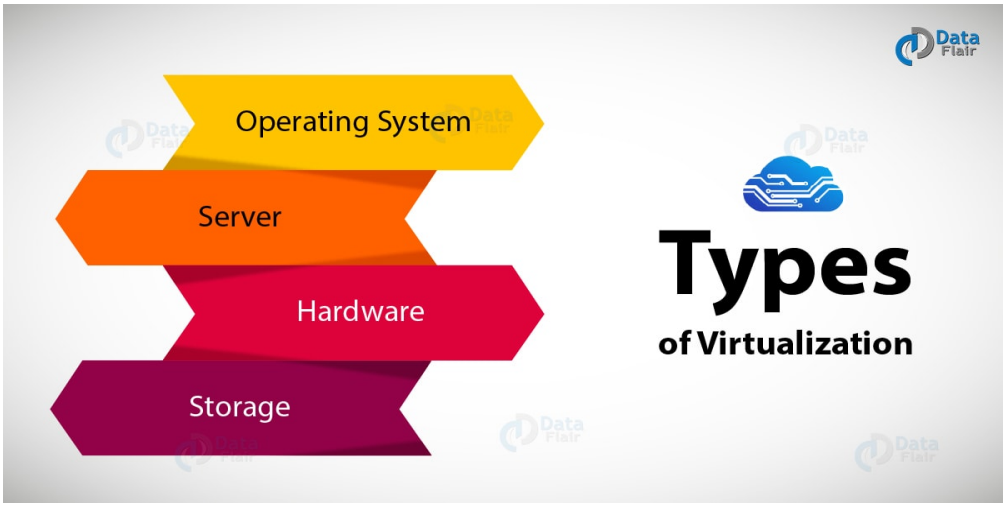


Figure Source Url

The virtualization of the cloud has been categorized into four different types based on their characteristics. These are:

1. Hardware Virtualization
    1. Full Virtualization
    2. Emulation Virtualization
    3. Para-virtualization
2. OS Virtualization
3. Server Virtualization
4. Storage Virtualization

Hardware Virtualization

It is the abstraction of computing resources from the software that uses cloud resources. It involves embedding virtual machine software into the server's hardware components. That software is called the hypervisor. The hypervisor manages the shared physical hardware resources between the guest OS & the host OS. The abstracted hardware is represented as actual hardware. Virtualization means abstraction & hardware virtualization is achieved by abstracting the physical hardware part using Virtual Machine Monitor (VMM) or hypervisor. Hypervisors rely on command set extensions in the processors to accelerate common virtualization activities for boosting the performance. The term hardware virtualization is used when VMM or virtual machine software or any hypervisor gets directly installed on the hardware system. The primary task of the hypervisor is to process monitoring, memory & hardware controlling. After hardware virtualization is done, different operating systems can be installed, and various applications can run on it. Hardware virtualization, when done for server platforms, is also called server virtualization.

Subtypes:

- **Full Virtualization** – *Guest software does not require any modifications since the underlying hardware is fully simulated.*
- **Emulation Virtualization** – *The virtual machine simulates the hardware and becomes independent of it. The guest operating system does not require any modifications.*

- **Para Virtualization** – *the hardware is not simulated and the guest software runs their own isolated domains.*

OS Virtualization

In operating system virtualization in Cloud Computing, the virtual machine software installs in the operating system of the host rather than directly on the hardware system. The most important use of operating system virtualization is for testing the application on different platforms or operating systems. Here, the software is present in the hardware, which allows different applications to run.

Server Virtualization

In server virtualization in Cloud Computing, the software directly installed on the server system and used for a single physical server can divide into many servers on the demand basis and balance the load. It can be also stated that the server virtualization is masking of the server resources which consists of number and identity. With the help of software, the server administrator divides one physical server into multiple servers.

It is the division of physical servers into several virtual servers and this division is mainly done to improvise the utility of server resources. In other words it is the masking of resources that are located in a server which includes the number & identity of processors, physical servers & the operating system. This division of one physical server into multiple isolated virtual servers is done by the server administrator using software. The virtual environment is sometimes called the virtual private-servers.

In this process, the server resources are kept hidden from the user. This partitioning of physical servers into several virtual environments; result in the dedication of one server to perform a single application or task.

Storage Virtualization

In storage virtualization in Cloud Computing, a grouping is done of physical storage which is from multiple network storage devices. This is done so it looks like a single storage

device. It can be implemented with the help of software applications and storage virtualization is done for the backup and recovery process. It is a sharing of the physical storage from multiple storage devices.

The crucial backbone technology for cloud computing is virtualization. The following points are worth noting.

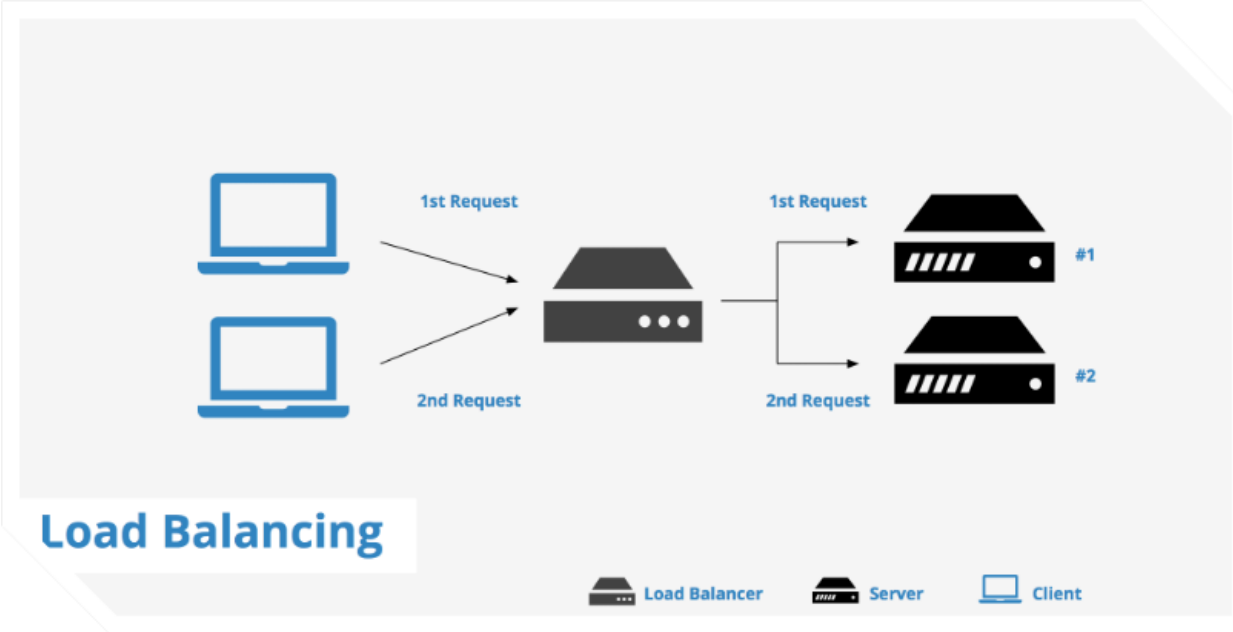
I. The existence of a cloud without virtualization, although possible, would be very inefficient & difficult.
II. The idea of infinite availability, Pay for as much you use, & Pay as you grow require a very flexible & efficient back-end which is readily available in virtualized environments and machines.

Load balancing in Cloud Computing

Load balancing in Cloud computing is one of the most challenging and useful research for distributing the tasks among the virtual machines at the Data Centers*.Load balancing technique balances the load of a machine by dynamically distributing the local workload equally to all nodes in the cloud to prevent the problem of over and underutilized nodes. Sharing loads over the distributed system ensures better utilization of resources and increases the throughput which leads to achieving user's Quality of service (QOS).Failing of one or more components of any services, Load Balancing ensures continuation of services using fair-over technique that means provisioning as well as de-provisioning related to application instances without failure. It makes sure that every computing resource is distributed fairly and efficiently. Hence, Load balancing aims to achieve better performance of systems with reasonable cost. In brief the goal of load balancing is increasing system stability by building fault tolerance systems without affecting the basic computational task. The other benefits of load balancing includes: failover- in case of failure it may forward network traffic to other servers; scalability – to increase computational capacity without

affecting other network/system components customers can quickly add servers under the load balancer.

**Datacenter is a repository of servers hosting different applications. End users who require services have to connect with the datacenter to subscribe to the services.*



Load Balancing Algorithms

Load balancing algorithms are used to distribute new requests of users in a data center between the Virtual Machines, in order to guarantee an equal number of tasks allotted to each machine.

Load balancing algorithms are basically two types :

I. static
II. dynamic.

Static algorithms

In a static algorithm, the decisions are made at compile time where requirements are initially estimated. Prior knowledge about the system is required without depending on the systems current state. Static algorithms work properly only when effective workloads are low. In addition, static approaches are defined usually in the design or implementation of systems. Their main drawback is that the current state of the system is not considered when making the decisions and therefore it is not a suitable approach in systems such as distributed systems where most states of the system change dynamically. Some of these algorithms are as follows :

- ***Round robin algorithm***

This task allocation order is based on the FCFS technique. Requests are distributed among VMs in turns using a data center controller regardless of the different allocated tasks characteristics . This algorithm is known for its fair allocation of the VMs to all nodes , but it doesn't have control over the workload distribution.

- ***Weighted round robin algorithm***

This algorithm is similar to the 'round robin' in a sense that the manner by which requests are assigned to the nodes is still cyclical, except that the node with the higher specifications will be given a greater number of requests

- ***Improved weighted round robin***

This algorithm is based on the 'weighted round robin'. It is similar in regard to the cyclic tasks allocation, except that it considers the priority and the length of the task to choose the most suitable VM .

And many more.....

A *general disadvantage of all static schemes is that the final selection of a host for process allocation is made when the process is created and cannot be changed during process execution to make changes in the system load.*

Dynamic algorithms

In a dynamic algorithm, every resource is provisioned or deprovisioned by load balancer at run time. It uses the system's current state information to make its decisions. These load balancing algorithms are adaptive and adapt their activities by dynamically changing their parameters, policies and system state. Dynamic techniques are highly successful for load balancing.

- ***Least connections***

This algorithm is based on the selection of the service with the least number of active connections to ensure that the load of the active requests is balanced on the services. This method is the default method because it provides the best performance .

- ***Weighted least connections***

This algorithm introduces a "weight" component based on the respective capacities of each server. Just like the 'weighted round robin', each server's "weight" must be specified beforehand. A load balancer that implements the 'weighted least connections' algorithm considers two things: the weight capacities of each server and the current number of clients currently connected to each server .

And many more……..

Metaheuristic algorithms

Metaheuristic algorithms have been an interesting area for researchers, scientists and academicians due to their specific and significant characteristics and capabilities. Their ability to solve and give near optimal solutions to the problems of versatile domains without in-depth details and definition of the problems, provides an edge over traditional techniques. Most of the metaheuristic algorithms are inspired by some real world phenomenon, generally a natural method of optimization. Over the last few decades a number of metaheuristic algorithms have been introduced and applied on various problems of different domains.

Some metaheuristic algorithms

- ***Ant colony optimization*** ***

This algorithm, as its name states, is based on the behavior of ants to detect the location of underloaded or overloaded nodes. It then updates the resources utilization table .This algorithm is known for its scalability but has low throughput.

- ***Honey Bee Algorithm*** ****

This approach uses the artificial bee colony algorithm with the feature of minimal migration time.This algorithm aims at realizing a well-balanced load among all VMs in order to maximize the system efficiency. The suggested algorithm equally balances the tasks priorities on machines so that the time spent in the waiting queue may be minimized

Metrics for Load Balancing

1. ***Throughput****- Effective load balancing strategy should increase systems throughput. Throughput is the number of tasks whose execution is completed to the total number of tasks within a time frame.*
2. ***Fault tolerance*** *- It means recovery from failure.This parameter shows if an algorithm is able to tolerate tortuous faults or not. It enables an algorithm to continue operating properly in the event of some failure. If the performance of the algorithm decreases, the decrease is proportional to the seriousness of the failure. Even a small failure can cause total failure in load balancing .A good load balancing strategy must handle the fault tolerance issues.*
3. ***Migration time****- The time that is required migrating a job or resources from one node to another. It decides whether to create it locally or create it on a remote processing element. The algorithm is capable of deciding whether to make changes of load distribution during execution of the process or not. It should be minimized for better performance of the system.*
4. ***Response Time*** *- The time taken by the load balancing algorithm to serve the first response of the job in a system. Load balancing algorithms should minimize the response time for better performance.*

5. ***Scalability*** *– It is the ability of the load balancing strategy that can balance any amount of load at any point of time with a finite amount of resources. The metric should be improved to increase throughput and response time.*

Goals of Load Balancing

1. *To improve the performance of resource utilization.*
2. *To improve the performance substantially at lower cost.*
3. *To have a backup plan in case the system fails even partially and it should be flexible enough to handle this situation.*
4. *To maintain the system stability by increasing the scalability.*
5. *Prioritization of the resources and as well as jobs that need to be done. So that high priority job gets a better chance to execute.*

HONEYBEE'S BEHAVIOUR

The artificial bee colony algorithm(ABC), is an optimization algorithm which is based on the smart foraging behavior. This Meta heuristic is inspired by the smart foraging behavior of honey bee swarm. The ABC contains three groups of bees [1]. They are

- **Scout bees** - *Simple meaning of scout is a soldier sent to gather information. The bee who carries out random search is identified as a scout. Upon finding one they returned to the beehive and informed forager bees.*
- **Forager bees** - *The bee which is going to the food source which is visited by scout bees previously is forager bee.*
- **Onlooker bees** - *Onlooker means somebody who watches an event without participating in it. The bee waiting on the dancing area is an onlooker bee.*

Scout bees are sent to search for suitable sources of food, when one is found, they return to the hive to advertise this using a display dance known as a "waggle dance". This dance shows the appropriateness of the food source. Forager bees then trail the scout bees back to the discovered food source and begin to collect it .The remaining amount of food

available is reflected in their waggle dances, on the basis of how many bees have been returned to the hive which allows more bees to be sent to a source, or exploited sources to be discarded.

Related Works

Load balancing performs the operation of balancing tasks across the under loaded VM and over loaded VM.

In [6],"***A Novel Honey Bee Inspired Algorithm for Dynamic Load Balancing In Cloud Environment***",the authors have discussed different honey bee inspired algorithms. Here the proposed algorithm uses the concept of pareto dominance"s weighted sum approach for selecting optimal VM and also works with preemptive task scheduling.

In [7],"***An Intensification of Honey Bee Foraging Load Balancing Algorithm in Cloud Computing***", the authors have discussed the performance of the honeybee foraging algorithm based on throughput calculated for each virtual machine.

In [8],"***Honey Bee Based Load Balancing in Cloud Computing***",the authors have designed load balancer which chooses most efficient node based on two conditions:number of tasks currently processing by this VM is less than number of tasks currently processing by other VMs and the deviation of this VM processing time from average processing time of all VMs is less than a threshold value.

In [1],"***Balancing Techniques: Need, Objectives and Major Challenges in Cloud Computing***" ,the authors have done a comparative analysis between dynamic and static load balancing scheme based on different criteria including performance, scalability, throughput, resource utilization, fault tolerance, response time etc.

In [2],"***Honey Bee Behavior Based Load Balancing of Tasks in Cloud Computing***" ,the authors have proposed a flow chart based on the nature exhibited by a foraging honeybee. Studies have proven that load balancing techniques inspired from honey bee nature upgrade throughput of processing. Response time of a VM is decreased by implementing a priority based balancing approach. They have also presented a comparison of their algorithm with other existing ones and have succeeded in proposing an algorithm that performs well without any considerable overhead.

In [3],"**A *Survey of Various Load Balancing Techniques And Challenges In Cloud Computing***",the authors have discussed various load balancing techniques for cloud computing had surveyed. They had also discussed cloud virtualization and required a qualitative matrix for load balancing. By referring to this paper we were able to identify the existing challenges in load balancing.

In [4] ,"**A *Comparative Study of Load Balancing Algorithms in Cloud Computing Environment***",The authors have reviewed different load balancing strategies, discussing their merits and demerits. While static load balancing offers simple and uncomplicated simulation and supervision of the environment as opposed to dynamic load balancing, it fails to offer the best solution when it comes to heterogeneous cloud environments. Dynamic load balancing on the other hand is better suited in heterogeneous cloud environments. If the algorithm is implemented and distributed in nature it succeeds in achieving fault tolerance, but with a greater rate of replication. To achieve a more balanced state, hierarchical load balancing algorithms can be employed where the load gets distributed to different hierarchical levels and nodes at hierarchical level request services from the nodes at lower level in a fair and balanced manner. Although implementation of dynamic load balancing techniques in a distributed or hierarchical cloud environment provides remarkable performance, modeling the dependencies between tasks using workflows can further boost the performance of the cloud.

In [5], "***Honey Bee Behavior Load Balancing of Tasks in Cloud***",the authors have focused on achieving minimum task completion time and better resource utilization.This they achieved by reallocating prioritized task from one machine to another having fewer number of prioritized task with minimum completion time. Besides considering the time required to complete a task, their study also focused on task's priority at the time of submission of execution, and also adjusting the load of dependent tasks in a preemptive manner.

Proposed Work

As was stated earlier, Honey bee foraging algorithm is an artificial manifestation of the way in which bees find and gather their food. Food is searched by the Forager bees and they announce it to other bees when they get it. They announce it by doing waggle dance which shows the availability of metadata food. Then the scout bees follow the forager bees and collect honey from the source. After returning, they do a waggle dance to indicate the food left so that more honey can be consumed. In our curated private Cloud Environment, the

servers are made into a group of virtual servers and each virtual server maintains a process queue. Processing Time computation causes additional overhead for throughput calculation , which results in deterioration of overall throughput itself. First we will be calculating the throughput and based on the highest throughput value we will be selecting that particular VM which will be assigned the next task to process.

Algorithm

We will be calculating throughput at the time of creating the VM. Step

Step 1: *Start Step*

Step 2: *Set number of tasks Step*

Step 3: *Set number of VM Step*

Step 4: *Calculate throughput.*

Step 5: *Initially, set the load on each VM to null Step*

Step 6: *Send first task to the VM having high throughput Step*

Step 7: *Check current VM load > threshold value*

Step 8: *If yes, select the VM having high throughput and check the load < threshold value, if yes, assign to that VM If no, select another VM with next highest throughput and load < threshold value*

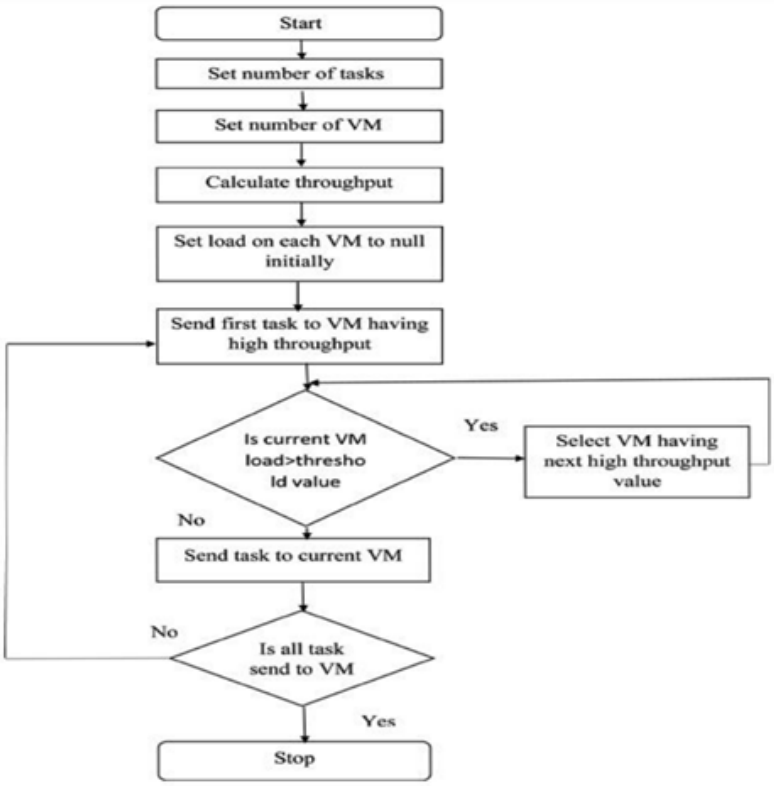
Step 9: *If no, send the task to the current VM*

Step 10: *If all task are send to the VM go to step 12*

Step 11: *Else go to step 6 Step*

Step 12: *Stop*

Flowchart



Flowchart of modified honey bee foraging algorithm

Conclusion and Future Prospects:

Here the proposed algorithm follows the foraging behaviour of honey bees for allocating VMs to the request processing and uses preemptive task scheduling in accordance with the server having the best throughput to tasks/requests are assigned while other tasks are assigned to the next highest throughput server. This way multiple tasks are done concurrently which gives the Whole system efficiency and performance which thereby increases productivity.

Here tasks are considered as independent tasks. In future work this honey bee inspired load balancing algorithm can be extended for dependent tasks .Calculation of assigning priority of task and finding optimal VM can be improved by considering QOS parameters as well which we have never considered.

References

[1] Nitin Kumar Mishra, Nishchol Mishra, Load Balancing Techniques: Need, Objectives and Major Challenges in Cloud Computing-A Systematic Review, International Journal of Computer Applications 131(18) (2015).

[2] Harshit Gupta, Kalicharan Sahu, Honey Bee Behavior Based Load Balancing of Tasks in Cloud Computing, International Journal of Science and Research (IJSR) (2012).

[3] Tushar Desai, Jignesh Prajapati, A Survey of Various Load Balancing Techniques And Challenges In Cloud Computing, International Journal of Scientific & Technology Research 2(11) (2013).

[4] Mayanka Katyal, Atul Mishra, A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment, arXiv preprint arXiv:1403.6918 (2014).

[5] Khushbu Zalavadiya, Dinesh Vaghela, Honey Bee Behavior Load Balancing of Tasks in Cloud Computing, International Journal of Computer Applications 139(1) (2016).

[6] Chandrakanta Korat, Piyush Gohel,A Novel Honey Bee Inspired Algorithm for Dynamic Load Balancing In Cloud Environment,International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 4(8), August 2015

[7] V.V. Bhavya,K.P. Rejina, A.S. Mahesh,An Intensification of Honey Bee Foraging Load Balancing Algorithm in Cloud Computing ,International Journal of Pure and Applied Mathematics ,Volume 114 No. 11 2017, 127-136

[8] Walaa Hashem, Heba Nashaat, and Rawya Rizk,Honey Bee Based Load Balancing in Cloud Computing, KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS VOL.11,NO. 12, Dec. 2017