IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 24, Issue 4, Ser. I (Jul. –Aug. 2022), PP 22-55
www.iosrjournals.org

Energy Efficiency in Cloud Computing Based On Optimization and Predict Analysis

Ravi Pratap, Sandeep Kumar
Department of Computer Science, BBAU Lucknow
Department of Computer Science, BBAU Lucknow

Abstract: *Cloud computing is an exceptionally versatile and effective foundation for running undertaking and web applications. Vitality utilization use and simultaneous impacts on condition are the dynamic difficulties with respect to distributed computing. We examined past investigates dependent on vitality productivity approaches and recovered the conditions to advance green distributed computing models. Server farms are considered as the foundation of cloud framework. Various organizations are putting resources into building up vast server farms to oblige diverse distributed computing administrations. These server farms assimilate tremendous measure of vitality. They are additionally mind boggling in the foundation. Throughout the years, control utilization has risen as an essential factor for estimating registering assets consumption. In this paper we will examine every attainable circle in a perfect cloud framework which are in charge of gigantic measure of vitality utilization. Here, we will likewise examine the systems by which vitality usage can be diminished without bargaining Quality of Services and by and large execution. So as to make these server farms more vitality productive, numerous examinations have been proposed. These looks into depend on advancements, for example, combination and virtualization. Diminishing discharges of carbon dioxide and vitality utilization set up new difficulties. The exploration works for green server farms are gotten from the difficulties. There is a present interest of coordinated vitality proficient cloud system for setting up server farms that limits the effects on condition and decreases CO2 discharges. The structure should consolidate a green IT design with various systems and exercises.*

Key-words*: Cloud Computing, Predictive Analysis, Convex Optimization, and Energy Efficiency.*

Date of Submission: 28-06-2022 Date of Acceptance: 09-07-2022

I. Introduction

Cloud computing is another utilization and virtualization show for the mind-boggling expense registering frameworks and is an electronic IT arrangement. Cloud gives reasonable, on-request administration, organize get to, asset sharing and estimated administration in very adjustable way with negligible administration exertion. The use of ease processing gadgets, superior system assets, colossal capacity limit, semantic web innovation, and so on., have helped in the quick development of cloud innovation. A cloud foundation for the most part embodies each one of those current innovations in a web administration based model to offer improved versatility and on interest accessibility. The fast organization display, low start-up venture, pay-as-you-go conspire, multi-occupant sharing of assets are altogether included properties of cloud innovation because of which real businesses watch out for virtualization for their undertaking applications.

Cloud applications are conveyed in remote server farms (DCs) where high limit servers and capacity frameworks are found. A quick development of interest for cloud based administrations results into foundation of gigantic server farms expending high measure of electrical power. Vitality proficient model is required for complete framework to lessen useful expenses while keeping up fundamental Quality of Service. Vitality enhancement can be accomplished by consolidating assets according to the present usage, productive virtual system topologies and warm status of figuring equipment. And nodes. On the other hand, the primary motivation of cloud computing is related to its flexibility of resources.

What is Cloud Computing?

Cloud computing can be defined as a "network based environment its aim is sharing Computations and resources in order to deliver service over internet". Saving in the energy budget of data-center with providing pay-as-you-use service and access the resources as multiple Virtual Machines are hosted on the same physical server on consumer request with minimum management risk.

(User pays only for what they are using; not pay for local resources like storage or infrastructure.)

Fig (1) Cloud Computing

National Institute of Standards and Technology (NIST) defines Cloud computing as follows:

"Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This Cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models."

Cloud computing as a new paradigm has changed the way in which computing is done. The flexibility and ease to acquire and release resources on-demand is truly great for the customers. Because cloud computing is such a large industry now, it makes sense to focus our energy efficiency efforts on this particular kind of computing. The requirements and peculiarities are different enough from request-parallel workloads to allow full-system power Modesto be applied. The characteristics of Clouds include on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. The available service models are classified as SaaS (Software-as-a-Service), PaaS (Platform-as-a-Service), and IaaS (Infrastructure-as-a-Service). The deployment models is categorized into public, private, community, and hybrid Clouds.

Cloud Computing characteristics:-

The characteristics of cloud computing are:-

- Reduction of Cost There are a number of reasons to attribute Cloud technology with lower costs. The billing model is pay as per usage; the infrastructure is not purchased thus lowering maintenance. Initial expense and recurring expenses are much lower than traditional computing.
- Elasticity Services stored by cloud are rapidly provisioned, and in few cases by itself, rapidly released to quickly scale in. To consumer, the capabilities available for provisioning appears to be endless and could be purchased whenever required.
- Security and Availability The cloud should authorize the data access by the end users. However a fear of security always there with the end users. There quests have to be fulfilled in every case, the data and infrastructure should always available.
- Flexibility Cloud computing mainly stress on deployment of applications in market as quickly possible, by using the most appropriate building blocks necessary for deployment.
- Geographical independence A user can access the data shared on cloud from any location across the globe

Various services associated with Cloud computing are:

i. **Infrastructure as a service (IaaS):** In this vendor's of cloud computing share their resources with clients such that clients have to pay as per use only. It allows clients to manage the operating systems, storage, applications, and network connectivity.

ii. **Software as a service (SaaS):** It works upon pay-per-use costing model. In SaaS software that are available on the cloud server are provided as a services to the consumers on the basis of their requirement.

iii. **Platform as a service (PaaS**): PaaS allows platform access for clients with the help of which they can put their own applications and soft wares on the cloud.

iv. **Communication as a service (CaaS):** In the CaaS of cloud computing services, CaaS falls between Infrastructure as a Service (IaaS) and Platform as a Service (PaaS).

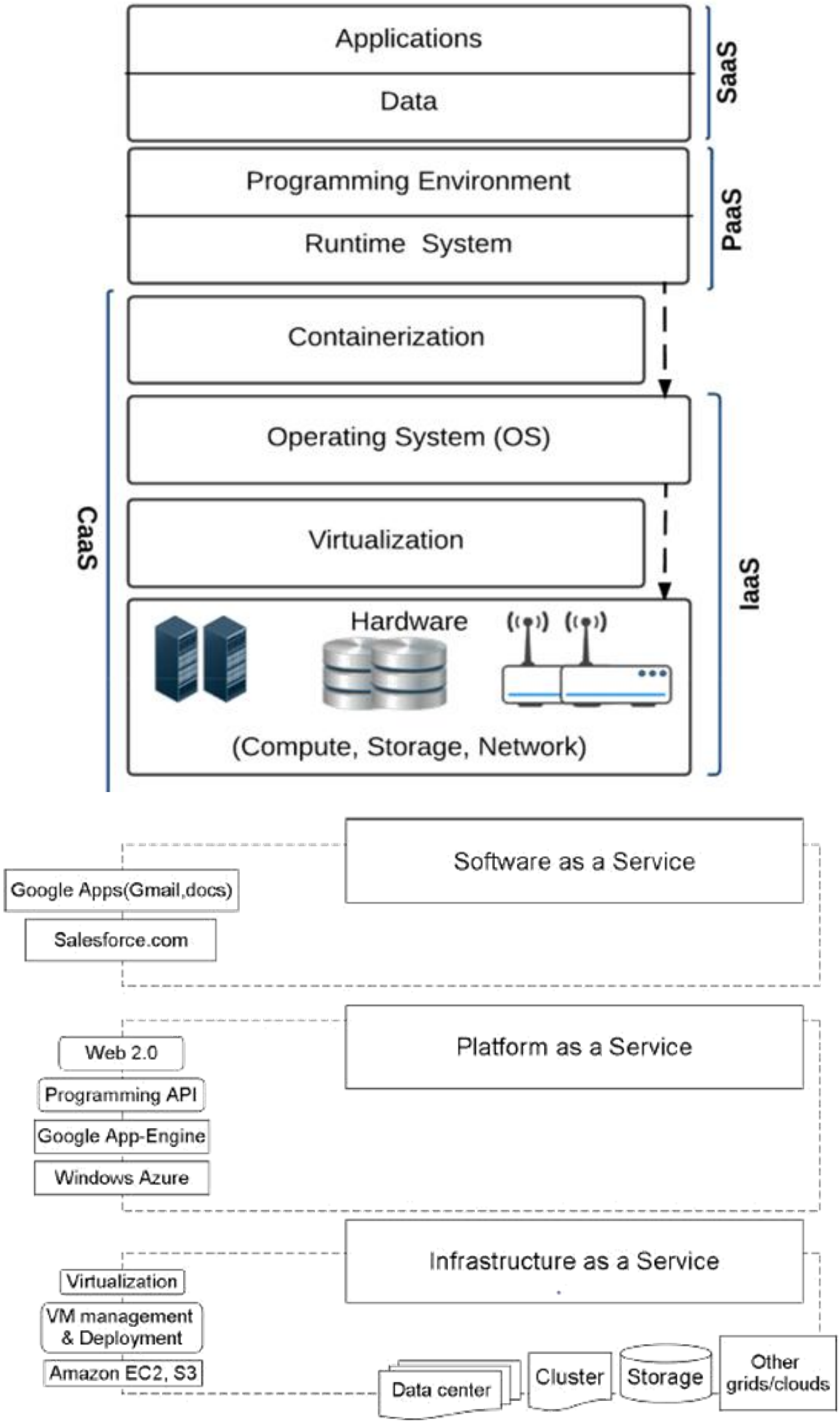


Figure (1): Cloud Service Stack

In Cloud computing important objective's for research community is Focusing on Various techniques.

- Cloud computing Deployment Models
- Energy Efficiency of cloud-computing.
- Energy-aware Data center's

Cloud Computing Deployment Model

The deployment model are shown in the diagram and explained as follows:

Public Cloud:-

A public cloud network enables users to distribute and access data from anywhere at any given point in time. This means that public cloud computing systems are incredibly accessible and can be shared with third parties. Based on the standard cloud computing model, in a public cloud the service provider makes its applications, storage or other resources, available to the general public. Examples of the public cloud include Google App Engine. The main benefits of a public cloud service are: easy and inexpensive to setup, scalability, and a pay per what you use model (no wasted resources).

Private Cloud:-

Availability and distribution mediums in a private cloud network are limited only for authorized users from behind a firewall. This form of cloud computing is specially designed for companies that do not want to distribute their internal work information to third parties. Nonetheless, these outside users can still access or distribute data provided they are authorized by the main client to access. Private cloud computing networks are much safer to use than public ones since they require all users to be authorized.

Hybrid Cloud

Hybrid cloud is developed with both public and private cloud characteristics. While public and private cloud systems are more prevalent, hybrid types have been growing in demand. Hybrid cloud systems occur when an organization provides some cloud services in-house and has others provided externally. The advantage to this approach is that companies are able to host external data on site with an external provider, while maintaining control over internal customer data.

Cloud Stakeholders

To know why cloud computing is used let's _first concentrate on who use it. And then we would discuss what advantages they get using cloud. There are three types of stakeholders cloud providers, cloud users and the end users [Figure 2]. Cloud providers provide cloud services to the cloud users. These cloud services are of the form of utility computing i.e. the cloud users uses these services pay-as-you-go model. The cloud users develop their product using these services and deliver the product to the end users.

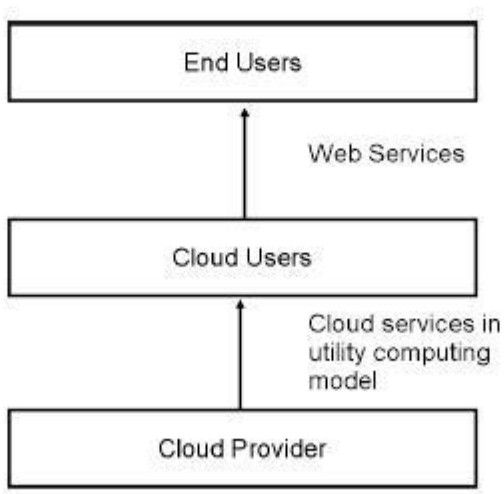


Figure 2: Interconnection between cloud stakeholders

II. Advantages of using Cloud

The advantages for using cloud services can be of technical, architectural, business etc.

1. **Cloud Providers' point of view**

(a) Most of the data centers today are under-utilized. They are mostly 15% utilized. These data centers need spare capacity just to cope with the huge spikes that sometimes get in the server usage. Large companies having those data centers can easily rent those computing power to other organizations and get profit out of it and also make the resources needed for running data center (like power) utilized properly.

(b) Companies having large data centers have already deployed the resources and to provide cloud services they would need very little investment and the cost would be incremental.

2. Cloud Users' point of view

(a) Cloud users need not to take care about the hardware and software they use and also they don't have to be worried about maintenance. The users are no longer tied to some one additional system.

(b) Virtualization technology gives the illusion to the users that they are having all the resources available.

(c) Cloud users can use the resources on demand basis and pay as much as they use. So the users can plan well for reducing their usage to minimize their expenditure.

(d) Scalability is one of the major advantages to cloud users. Scalability is provided dynamically to the users. Users get as much resources as they need. Thus this model perfectly it's in the management of rare spikes in the demand.

Motivation towards Cloud in recent time

Cloud computing is not a new idea but it is an evolution of some old paradigm of distributed computing. The advent of the enthusiasm about cloud computing in recent past is due to some recent technology trend and business models.

1. High demand of interactive applications {Applications with real time response and with capability of providing information either by other users or by nonhuman sensors gaining more and more popularity today. These are generally attracted to cloud not only because of high availability but also because these services are generally data intensive and require analysing data across different sources.
2. Parallel batch processing {Cloud inherently supports batch-processing and analysing tera-bytes of data very efficiently. Programming models like Google's map-reduce and Yahoo!'s open source counterpart Hadoop can be used to do these hiding operational complexity of parallel processing of hundreds of cloud computing servers.
3. New trend in business world and scientific community { In recent times the business enterprises are interested in discovering customer's needs, buying patterns, supply chains to take top management decisions. These require analysis of very large amount of online data. This can be done with the help of cloud very easily. Yahoo! Homepage is a very good example of such thing. In the homepage they show the hottest news in the country. And according to the users' interest they change the ads and other sections in the page.} Other than these many scientific experiments need very time consuming data processing.
4. Extensive desktop application { Some desktop applications like Matlab, Mathematica Are becoming so compute intensive that a single desktop machine is no longer enough to run them. } So they are developed to be capable of using cloud computing to perform extensive evaluations.

Cloud Architecture

The cloud providers actually have the physical data centers to provide virtualized services to their users through Internet. The cloud providers often provide separation between application and data. This scenario is shown in the Figure (3). The underlying physical machines are generally organized in grids and they are usually geographically distributed. Virtualization plays an important role in the cloud scenario. The data center hosts provide the physical hardware on which virtual machines resides. User potentially can use any OS supported by the virtual machines used.

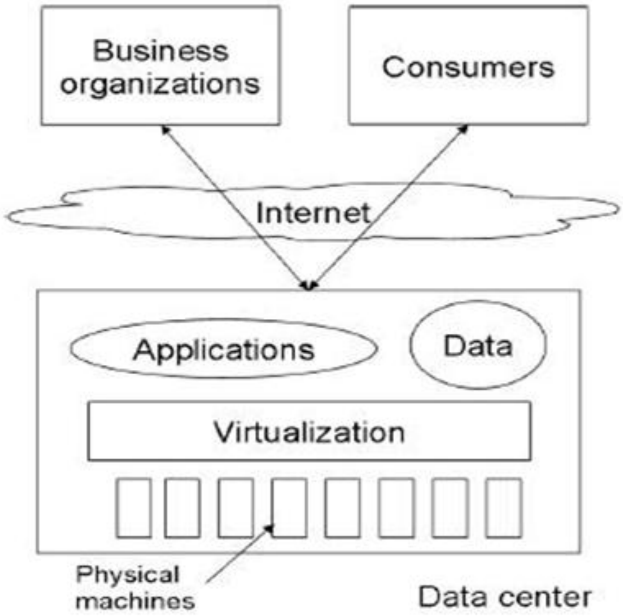


Figure 3: Basic Cloud Computing Architecture

- Operating systems are designed for specific hardware and software. It results in the lack of portability of operating system and software from one machine to another machine which uses different instruction set architecture. The concept of virtual machine solves this problem by acting as an interface between the hardware and the operating system called as system VMs. Another category of virtual machine is called process virtual machine which acts as an abstract layer between the operating system and applications. Virtualization can be very roughly said to be as software translating the hardware instructions generated by conventional software to the understandable format for the physical hardware. Virtualization also includes the mapping of virtual resources like registers and memory to real hardware resources. The underlying platform in virtualization is generally referred to as host and the software that runs in the VM environment is called as the guest.
- The Figure 3 shows very basics of virtualization. Here the virtualization layer covers the physical hardware. Operating System accesses physical hardware through virtualization layer. Applications can issue instruction by using OS interface as well as directly using virtualizing layer interface. This design enables the users to use applications not compatible with the operating system.
- Virtualization enables the migration of the virtual image from one physical machine to another and this feature is useful for cloud as by data locality lots of optimization is possible and also this feature is helpful for taking back up in different locations. This feature also enables the provider to shut down some of the data center physical machines to reduce power consumption.

Relation between Cloud Computing and Utility Computing

The cloud users enjoy utility computing model for interacting with cloud service providers. This Utility computing is essentially not same as cloud computing. Utility computing is the aggregation of computing resources, such as computation and storage, as a metered service similar to a traditional public utility like electricity, water or telephone network. This service might be provided by a dedicated computer cluster specially built for the purpose of being rented out, or even an under-utilized supercomputer. And cloud is one of such option of providing utility computing to the users.

III. Energy Efficiency In Cloud Infrastructures

Building an energy efficient cloud model does not indicate only energy efficient host machines. Other existing components of a complete cloud infrastructure should also be considered for energy aware applications. Several research works have been carried out to build energy efficient cloud components individually. In this section we will investigate the areas of a typical cloud setup that are responsible for considerable amount of power dissipation and we will consolidate the possible approaches to fix the issues considering energy consumption as a part of the cost functions to be applied.

A) Energy Efficient Hardware

One of the best approaches to reduce the power consumption at data centre and virtual machine level is usage of energy efficient hardware's at host side. New electronics materials like solid-state drives are more power efficient than common hard disk drives but that are costly. Some existing mechanisms for saving energy at hardware level are given below:

1. **SpeedStep®:** This is Intel's wireless technology to adjust CPU power dynamically based upon the performance demand. It works in five usage modes: standby mode, voice communication, multimedia, data communication, and multimedia and data communication.

2. **PowerNow! ™:** This is AMD‟s power saving technology which can manage power consumption instantly, on-the-fly by controlling voltage and frequency independently. It can be operated in three modes namely High Performance mode, Power saver mode and automatic mode.

3. **Cool„n‟Quite™:** AMD‟s Cool„n‟Quite™ technology controls the system fan, voltage and clock speed of the processor's cores based on the system temperature.

All these technologies are able to sense lack of machine interaction and then different hardware parts can incrementally be hibernated or put in sleep mode to save energy.

B) Energy Efficient Resource Scheduling

Several research works have been carried out on energy efficient resource scheduling in virtual machines and grid systems. Many aspects such as: power consumption, reliability, response time etc. need to be considered while designing VM schedulers.

1) Real time tasks scheduling in multiprocessor systems: Real time scheduling of tasks in a multiprocessor host machine has been a major issue to be considered. Energy aware, real-time task scheduling method in multiprocessor system supports Dynamic Voltage Scaling (DVS). It is a polynomial time heuristic approach for building probabilistic model of load balancing issue. Task scheduling can be of two types: i) partitioned scheduling – where assignment of each individual task to a particular processor is fixed and ii) non-partitioned (dynamic) scheduling – where a long term global scheduler assigns the jobs to processors on-the-fly from a ready queue. In a multiprocessor system, there is a set of pre-emptive, independent, real time tasks denoted as T = {T1, T2, ..., Tn} and a set of processor cores denoted as P = {P1, P2, ..., Pm} where each core has a finite range of discrete frequencies F = {f1, f2, ..., fk}.

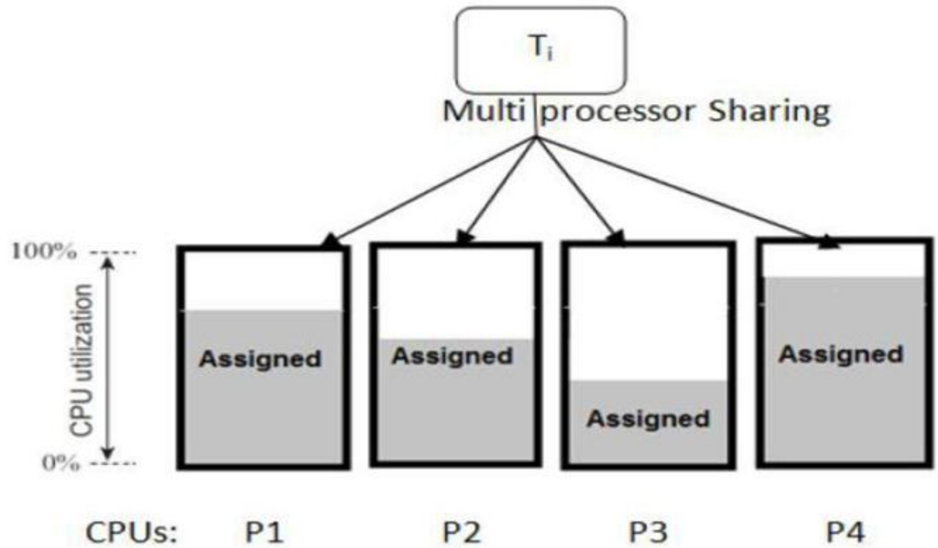


Figure 4: Multiprocessor sharing of a real time task

(1) A processing unit consumes both static power (constant) and dynamic power is during busy period and only static power when the unit is idle. Following Early Deadline First (EDF), an optimal scheduling algorithm, DVS aims to minimize only dynamic power (as static power dissipation is constant) for continuous, unbounded

frequencies and discrete, bounded frequencies of the system. New researches have proposed to implement bat intelligence (BI) along with Genetic Algorithm (GA) to solve energy-aware scheduling problem with improved QoS (Quality of Service) in multiprocessing environment. Thus power consumed by scheduling operation can be reduced substantially.

(2) Memory-aware Scheduling in multiprocessor systems: Optimization of scheduling algorithm helps to reduce the energy consumption by memory and also affects the efficiency of frequency scaling. In present multicore systems, cores on the chip share resources

Such as caches, DRAM etc. Tasks running on one core may harmfully affect the performance of tasks on other cores and hence it may even maliciously create a Denial of Service (DoS) attack on the same chip.

C) Energy efficient Network Infrastructure in cloud

Network in a cloud environment can be of two types - wireless network and wired network. According to ICT energy estimates the radio access network consumes a major part of the total energy in an infrastructure and the cost incurred on energy consumption is sometimes comparable with the total cost spent on personnel employed for network operations and maintenance. In order to achieve energy savings, use GAF and CEC protocols. These protocols identify redundant nodes and turn off them to conserve energy.

1) Geographic Adaptive Fidelity Protocol: In GAF protocol, equivalent nodes are found out by using their geographical information and then their radios are turned off which saves energy. However for communication between a pair of nodes, nodes which are equal may not be equal for communication between a different pair. This problem is addresses by dividing the whole network into virtual grids which have the property that all nodes in adjacent grids can communicate with each other. All nodes within a single grid are equivalent. Also the nodes in GAF protocol always switch among one of the three states – sleeping, discovery and active. Initially a node is in discovery state with its radio turned on and it exchanges messages with its neighbours. A node in active state and discovery state can switch to sleeping state whenever it finds an equivalent node which can perform routing.

2) Cluster Based Energy Conservation Protocol: One of the disadvantages with GAF protocol is that it needs global location information which may not be available every time. Also it is very conservative because it guesses its connectivity instead of directly measuring it which leads to less energy savings. Another protocol namely Cluster based Energy Conservation overcomes these disadvantages since it is independent of location information and finds network redundancy more accurately thus saving more energy. CEC operates in following steps-

- Determining Network Redundancy: It operates by organizing nodes into clusters which are overlapping. A cluster can be viewed as a circle around the node called cluster head. A cluster head is defined in such a way that it reaches each node in a cluster in just one hop. Another node called gateway node interconnects two or more clusters is a member of multiple clusters and overall network connectivity is ensured by it. Third type of node in a cluster is called ordinary node and is the redundant one because it neither the gateway node nor the cluster head.

- **Distributed Cluster formation:** A cluster selects a cluster-head and broadcast node by broadcasting discovery messages to its neighbour's. A discovery message is a combination of node ID, its cluster ID and its estimated lifetime. A node having the longest lifetime selects itself as the cluster head and this information is obtained by exchanging discovery messages. The node id of the cluster head becomes the cluster id of each node in a cluster. A primary gateway node is one that can hear multiple cluster heads and the one hearing the gateway node and cluster head is secondary gateway node. Some of the redundant gateway nodes among multiple gateway nodes between two clusters are suppressed by CEC to conserve energy. These redundant gateway nodes are selected according to certain rules. Primary gateway nodes have higher priority and are preferred over secondary because only one primary node can connect adjacent clusters. Finally redundant nodes are powered off after selecting cluster-heads and gateway-nodes thus conserving energy.

Energy savings in wired networks is not given as much importance as compared to energy savings in wireless networks. In fact, concepts applied for energy conservation in wireless networks cannot directly be applied in wired networks such as turning the nodes power off when not required to save energy because of high volumes and rate of data traffic and further nodes need to satisfy quality of constraint requirements. Also in a wired network, 60% of the energy consumed may be due to peripheral devices such as link drivers.

IV. Power Saving Techniques:-

Active vs. Idle Low-power Modes

In terms of energy efficiency, two fundamental device operating modes exist: active and idle. An active state allows the device toper form useful work, although possibly at a reduced rate, if it is a low-power active state. An idle state, on the other hand, characterizes a state in which the device cannot perform any work. Hence, three sensible states exist: full-power active, low-power active, and low-power idle. This is also the order from highest consuming state to lowest. Low-power idle states achieve higher energy savings compared to low-power active states due to the complete seizure of work. However, this comes at the expense of longer transition times until work can be resumed again. While the low-power active mode continues to get work done, although at a lower rate, an idle device must first be reactivated, incurring a delay.

To give an example, consider the suspend-to-memory functionality of modern servers. A suspend-to-memory transitions the server from an active full-power mode into an idle low-power mode. Only the network card and main memory continue to be supplied with power, while all other components are off. Because the CPU is off, the server cannot perform any work, e.g., answer requests coming in over the network. At the same time, the server's power consumption is reduced from more than 100Watt, depending on configuration, to 10 20Watt, a significant reduction. An active low-power mode example is the ability to dynamically adapt the processor's frequency. In times of low demand, the frequency can be varied in fixed, discrete steps. This technique is known as dynamic voltage frequency scaling (DVFS). However, DVFS only reduces the power consumption by tens of Watts. Also, the savings are limited to a single server component, the CPU, while other components are not designed to operate at different activity levels. The question of whether active or idle power modes can be applied boils down to how fast the system/component can transition between the full-power and low-power states. Suspending a Physical server takes on the order of a few seconds, as does resuming it. Because suspending cannot usually be interrupted, the minimum delay once a suspend was initiated is a few seconds. If the services running on that server have to guarantee latencies in the millisecond range, transitions on the order of multiple seconds preclude the use of these full-system idle states.

V. Migration Of The Application To The Cloud

Application migration is the process of transitioning all or parts of an enterprise data, applications and services from on-site premises and redeploying them on a new platform and infrastructure in the cloud. The migration process includes the staging of the new environment before the actual cutover for the application and requires Coordination between the IT teams in the time of migration.

A) Will all the application move to the cloud?

The answer is no. Some applications can be considered as good candidates to move fully or partially to the cloud platform, while other applications do not fit the cloud. To identify which applications can move to the cloud, applications are decomposed to several attributes form technical and business factors. The cloud is also decomposed into attributes to help make decision to run the application in the cloud or not.

1) Cloud computing is a Fit

Database Intensive Applications: It fits the cloud as the cloud computing and database are compatible. However, the cloud is more suitable for storing large volumes of unstructured data. On the other hand, when the data is accessed frequently in huge number of transactions in applications such as banking systems, then it can struggle the cloud. Such applications are governed by hardware limitations inside all IT environments, and thus there are no or little benefits from sending them to the cloud.

Network Intensive Applications: It is a good fit, but it is necessary to have an access to the fastest and highest quality network services in the cloud. Attributes such as network capacity, latency, redundancy and routing flexibility have to be taken into account when searching for the proper cloud provider.

Virtualized Applications: It is excellent fit. The cloud is completely compatible with such applications. The best way to ensure the cloud success is to optimize the applications in terms of CPU, storage, interface, and network performance.

2) Cloud Computing is not a Fit: Practically, the opposite characteristics for the applications or the classifications of applications mentioned in the previous section are not fit for cloud computing. Here more

detailed information will be added about the applications and the characteristics of the applications which are not recommended to move to the cloud.

Highly Secured Applications: when a high level of security is required, or the application has very sensitive data, then moving application to the cloud is not highly recommended. When company moves its IT service on the Cloud, they are no longer in control of the location of data.

Thus, the organization has to be careful when investigating the security level/controls that the service provider ensures.

When the control is very critical on the business some components in the applications are very critical to the business. If such components will not be moved to100% reliable cloud platform, then it could be endangering.

B) Attributes in the Migration Decision

The following shows some of these attributes and their relations to cloud computing based on some findings extracted mainly from the advantages and disadvantages of cloud, or in others word, what the cloud likes and what does not like.

1) Online Data Access:

This attribute affect the decision when the company wants to use the cloud storage services. If the application requires online access for data, then cloud would be an excellent choice as the company does not need to keep any data on-site, and thus it does not need to move huge bulk of data to their local servers.

2) Offline Data Access:

On the contrast to online access data, when an organization requires pure offline access of data, then it is better to keep data locally rather than put it on the cloud.

3) Structured Data:

Cloud supports structured data by services that provide database-like functionality. However, when the application is a database intensive with huge amount of transactions, this can cause some reliability issues. Structured data requires relational database such as SQL databases; SQL databases are difficult to scale, meaning that they are not natively suited to a cloud environment.

4) Unstructured Data:

It is more supported by cloud environment. An example of such data is a file system that stores media files. Unstructured data is non-relational or No SQL database; it is built to service heavy read and write operations and it is able to scale up and down easily, and thus it is more natively suited to run on the cloud.

5) Authorization/Authentication:

The enterprise authentication and authorization framework does not naturally extend into the cloud. The multi-tenancy architecture of the cloud force different mechanisms for authentication and authorization; cloud application services multiple tenants using one server and one application instance. Thus, there will be two kind of identification (authentication): tenant and user identification. Tenant can be identified by the domain name of the mail address for example, while the user identification can be done by using special user database for every tenant (company).

6) Replication:

Cloud is tightly related to replication technology. Replication can be performed for data or computing nodes. When the main replica crashes, one of the other replicas compensates the failed one instantaneously. Moreover, these replicas can be used in the load distribution so the application can handle the increasing demands.

7) Caching:

Cloud supports caching by adding another storage layer “distributed data cache” makes data access quick.

8) Pooling:

Cloud is a shared pool of computing resources such as network, server, storage, application, and service. Pooling is strongly correlated to the cloud environment.

9) Scale Out/Up:
Application migrated to the cloud can benefit from elastic nature of cloud computing, and therefore, scale-out/scale-up attributes are supported by the cloud with pay-as-you-use purchase model.

C) Migration Decision Calculation Process (Migration Algorithm)

In the preceding section, we have explained the role of each attribute and low level attribute in the migration decision to cloud computing environment, and how it can affect the organizations" decisions when they think to move their applications to the cloud. Every organization has their own combination of these attributes according to how these attributes are more critical for their businesses. For example, some organizations concerns in scalability more than affordability, while others are aware of the security and availability issues. As mentioned before, the first step in the concept is to tell the user (organization) how good or not its decision to move the application to the cloud. Thus, the system takes this preferable combination from different kind of attributes and makes some calculations (Migration algorithm) on them, and returns to the user the score of this combination and the recommended decision that organization should take. The following steps show in details the process to perform these calculations (Migration Algorithm).

1) Every attribute has been given a certain value (weight) represents the contribution of this attribute in cloud computing. In other words, how much the cloud computing supports the attribute. For example, online data access has been given a positive value as it is recommended by the cloud, while offline attribute has been given a negative value as it is not recommended moving the application that accesses its data offline to the cloud.

2) The weights mentioned in the first point are given for every application attribute. The idea behind these weights is to measure the impact for the attribute in the application migration decision as they will be used in the equation in the point 6; the attribute has a stronger relation with cloud will have higher weight and indeed will be weighted more in the equation. Not all weights are given positive values, some attributes are given negative value and these attributes can decrease the application score in the equation if they have been selected by the user, and consequently, reduce the chance to migrate the application to the cloud.

3) The organization can select any number or combination from these attributes even though that some attributes can have conflict contribution in the cloud decision. For instance, user can select both online and offline data access or structured and unstructured data storage which is a conflicted attributes combination.

4) To solve the issue of selecting some conflicts in the selected attributes, we added levels (importance levels) for every selected attribute such as low, high and critical. For example, if the user selects offline and online data attributes, online attribute can be more important for the application than offline, then the user is able to choose "High" or "Critical" levels for online and low level for offline.

5) In addition to the weights given to the attributes as numerical values, the levels (low, high, critical) are given numerical values as well. Low, high, and critical are assigned to 1, 2, and 3 respectively. Both weights and levels are used to calculate the application score which controls the migration decision.

6) From the selected combination from the attributes and their levels, the business logic (migration algorithm) calculates the following:

- The lowest value in the score scale (the lowest application score): it is the lowest score can be obtained from any selected combination the user can select from the available attributes. It is calculated as if the user selects the most negative combination of attributes and levels. To clarify, user only selects the attributes which have negative weights.

Moreover, user selects "Critical" as level for every negative attribute. Thus, any application score can be obtained later will not be less than this value. This lowest value represents will be the lowest point in the "ruler" that will measure the application score for the migration decision.

- The highest value in the score scale (the highest application score): it is the highest score can be obtained from any selected combination the user can select from the available attributes. It is calculated as if the user selects the most positive combination of attributes and levels. To clarify, user only selects the attributes which have positive weights.

Moreover, user selects "Critical" as level for every positive attribute. Thus, any application score can be obtained later will not be higher than this value. This lowest value represents will be the highest point in the "ruler" that will measure the application score for the migration decision.

- The real score value: it is the real score of the selected combination, and depending on this value the decision is taken to move the application to the cloud or not. The value is calculated according to the following formula:

$$\text{Application Score} = \Sigma Aw \times Lv$$

Where Aw = the weight of the attribute.
Lv = the value of the level

7) From the lowest the highest application scores calculated from the previous points, we define the scale "ruler" that can measure the score of the application to move to the cloud or not.

8) To add more flexibility to the score scale, we divided the resulted scale into five periods. Two periods in the

- The middle period starts with a certain value represent percentage from the lowest score. This percentage has been taken as 10%. The middle point ends with a value equal to 10% from the highest score. The idea behind this small period compared to the other four periods that we want to make a period where the user can decide between move the application to the cloud or not.
- The remaining scale is divided equally in both sides.

9) When the application's score comes in the first period that means that the application is not a cloud fit completely, while if the score comes in the last period that means that the application is strongly recommended to move to the cloud. The middle period gives the user the freedom to choose either to move the application or keep the thing as they are. The other two periods: one comes in the negative area and it does not support the migration decision, while the other one located in the positive area and it recommends the user to move the application to the cloud.

V. ENERGY CONSUMPTION ANALYSIS

1) To achieve energy efficiency at application level, SaaS providers should pay attention in deploying software on right kind of infrastructure which can execute the software most efficiently. For example, a simple task such as listening to music can consume significantly different amounts of energy on a variety of heterogeneous devices. As task have the same purpose on each device, the results show that the implementation of the task and the system upon which it is performed can have a dramatic impact on efficiency.

2) To calculate the amount of energy consumed by data centers, two metrics were established by Green Grid, an international consortium. The metrics are Power Usage Effectiveness (PUE) and Data Centre Infrastructure Efficiency (DCiE) as defined below:

PUE = Total Facility Power/IT Equipment Power
DCiE = 1/PUE = (IT Equipment Power/Total Facility Power) x 100%
The IT equipment power is the load delivered to all computing hardware resources, while the total facility power includes other energy facilities, specifically, the energy consumed by everything that supports IT equipment load.

3) In cloud infrastructure, a node refers to general multicore server along with its parallel processing units, network topology, power supply unit and storage capacity. The overall energy consumption of a cloud environment can be classified as follows:

ECloud = ENode + ESwitch + EStorage + EOhters

Consumption of energy in a cloud environment having n number of nodes and m number of switching elements can be expressed as:
ECloud = n (ECPU + EMemory + EDisk + EMainboard + ENIC) + m(EChassis + ELinecards + EPorts) + (ENASServer + EStorageController + EDiskArray) + EOthers

VI. Communication Networks

Communication systems have rarely been extensively considered in cloud computing research. Most of the cloud computing techniques evolved from the fields of cluster and grid computing which are both designed to execute large computationally intensive jobs, commonly referred as High-Performance Computing (HPC). However, cloud computing is fundamentally different: Clouds satisfy the computing and storage of millions of users at the same time, yet each individual user request is relatively small. These users commonly need merely to read an email, retrieve an HTML page, or watch an online video. Such tasks require only limited computation to be performed yet their performance is determined by the successful completion of the communication requests but communications involves more than just the data center network; the data path from the data center to the user also constitute an integral part for satisfying a communication request. Typical delays for processing users´ requests, such as search, social networks and video streaming, are less than a few milliseconds, and we sometimes even measured on the level of microsecond. Depending on the user location, these delays are as large as 100 milliseconds for intercontinental links and up to 200 milliseconds if satellite links are involved. As a result, a failure to consider the communication characteristics on an end-to-end basis can mislead the design and operational optimization of modern cloud computing systems.

Optimization of cloud computing systems and cloud applications will not only significantly reduce energy consumption inside data centers, but also globally, in the wide-area network. The World hosts around 1.5 billion Internet users [1] and 6 billion mobile phone users, and all of them are potential customers for cloud computing applications. On an average, there are 14 hops between a cloud provider and end users on the Internet. This means that there are 13 routers involved in forwarding the user traffic, each consuming from tens of watts to kilowatts. According to Nordman, Internet-connected equipment accounts for almost 10% of the total energy consumed in the United States. Obviously, optimization of the flow of communication between the data center providers and end users can make a significant difference. For example, a widespread adoption of the new Energy-Efficient Ethernet standard IEEE 802.3az can result in savings of 1 billion Euro.

At the cloud user end, energy is becoming an even greater concern: More and more cloud users use mobile equipment (smart phones, laptops, tablet PCs) to access cloud services. The only efficient way for these battery-powered devices to save power is to power off most of the main components, including the central processor, transceivers and memory, while also configuring sleeping cycles appropriately. The aim is to decrease request processing time so that user terminals will consume less battery power. Smaller volumes of traffic arranged in bursts will permit longer sleeping times for the transceivers, and faster replies to the cloud service requests will reduce the drain on batteries.

Shared Resources:

(Server's, Network, storage, Applications, Software etc.) These resources are configurable on user demand. Most of the business enterprisers and individual IT Companies are opting cloud service in order to share their business information .Some of existing cloud service provider are ,Google, Amazon, Microsoft's Windows Azure and IBM .

Energy-aware Data centers:-

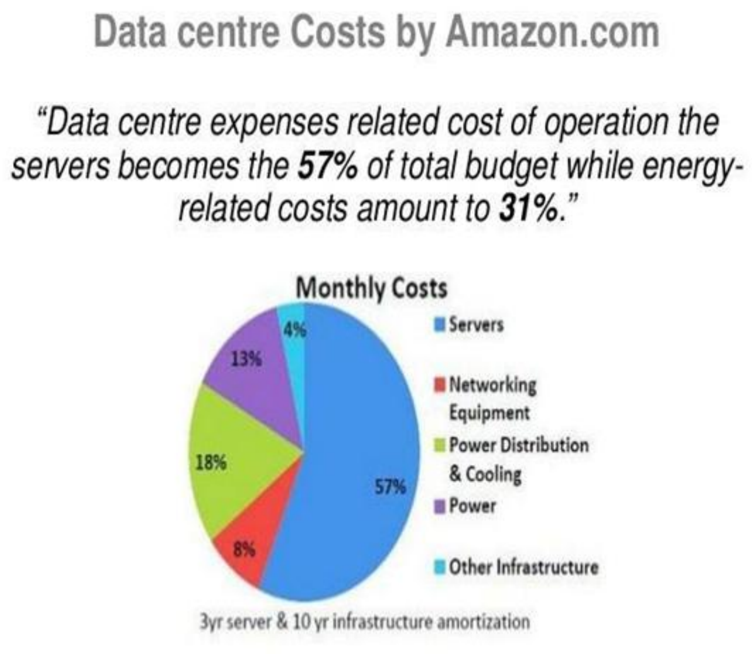


Figure . Energy distribution in Amazon DATA center

Data center Architecture

The Cloud Computing data centers are much flexible than the traditional data centers. In the traditional data centers all the servers are physically presented but now a day with developing trend of technology the virtualization is used to give best output to their customers. The data centers like Google, Facebook has a lot of VMs on one host to access the query of client in few seconds. A Data Center houses has hundreds to thousand number of storage units, servers which are connected with each other through various routers/switches, that are arranged to form one particular topology. The architecture of Data Center is shown in Figure 3 in which enough bandwidth is considered to overcome the problem of congestion4.

In the architecture of Data Center there are three main components broker, green service allocator and resources. Broker is the intimidator between the customer and resources. It takes cloudlets from the client and sends to the servers for processing.

Green service allocator layer: This layer provides the effective services to the customer by applying some particular actions on the system that the servers can allocate minimum machines to the cloudlets on request of broker. Green service allocator layer has six functions.

- Green negotiator: Green negotiators negotiate with the broker and client to set the SLA (Service Level Agreement) in between the service provider and the consumer according to the QoS requirement of the client and on specific pricing.
- Service analyzer: It checks the requirement of customer before accepting or rejecting the request on the basis of VM manager to maintain the minimum load and energy consumption.
- Consumer profiler: Consumer profiler analysis the background of consumer to give the priority according to the importance of consumer over the other consumers.
- Pricing: It calculates the charged of request that how it manages the demand and supply of computing resources according to the priority of the cloudlet.
- Monitor energy: It searches and evaluates the physical machine to power on or off.
- Schedule services: Service scheduler allocates the VM to the cloudlets and determine that when VM should add or remove according to the workload.
- Manage VMs: VM manager keeps record of available VM and also migrates VM on the other physical machines if needed.
- Accounting: It maintains the usage of machines to compute the usage cost.

The next component is resources in which physical machines are involved with the virtualization in it. Each physical machine has more than one VM to reach the accepted demand of customer. The VM are dynamically getting start and stop on the physical machine according to the work load. With the development of technology VM are migrating on the different physical machines to give low SLA violation and consume less energy. The number of parameters can be considered by the researcher to make the improvement in the service level of data center. Virtual Machines are separated on the basis of their use:

- A system virtual machine: It is the complete operating system which use then real hardware are not available and it is also known as hardware virtualization.

- A process virtual machine: It is used to run single program. These types of machines are built for providing portability and flexibility to the program and it is also called application virtual machines.

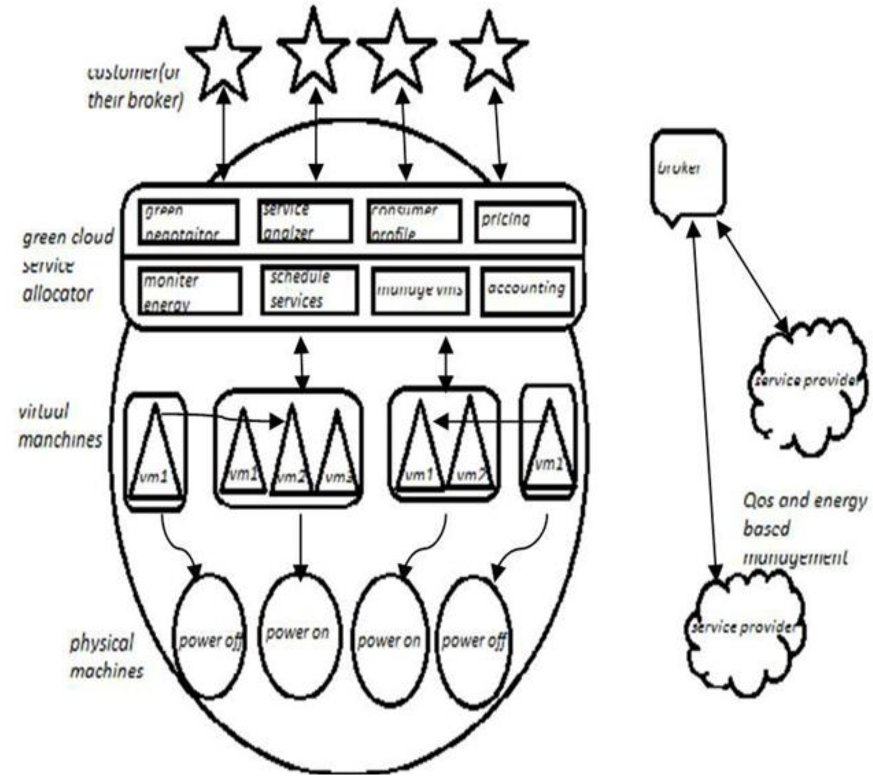


Figure Data Center Architecture

Energy Consumption:-

Since Cloud Computing came into existence are not perfect in terms of energy consumption. The consumption of large amount of energy has negative impact on our environment by releasing huge CO2 which arise

Greenhouse effect. Energy consumption includes many factors such as load balancing, power distribution, cooling, server and etc. Addition to this in 2007 one report is submitted to the US congress on “Server and data Centre energy consumption”. This report focus on energy efficiency by US data Centre was 61 billion kilowatts-hour in 2006 totaling was $4.5 billon5. The higher power consumption require cooling system that cost in the range of $2 to $5 million per year. The main two factors which help data centre to consume less energy are:

1. Shutting it down,
2. Scaling down its performance
3. When Cloud Computing came into existence its main focus was to make it a huge data centre for high performance computing and making profit from it by get paid what we used, but with the passage of time, it became a model of computing facility for the dynamic provisioning
4. Nowadays VMs migration and consolidation algorithms are based on energy consumption model mostly with single system resource constraint, i.e. CPU. These algorithms may not assume the impact of other resources. Some researchers believe that the energy consumption of whole server varies mostly linearly with the CPU utilization. However, it still consumes more than 70% of its peak energy, even if a server is completely idle. Day to day usage of computing services leads to energy consumption. The energy consumption has huge impact on the environment with the dissipation of CO2, which increases greenhouse effect
5. When cloud perform operation, cloud use data Centre to store data, process stored data with the help of servers and data gets it transferred over the internet.

It is estimated that approximately 10% of the world total energy is consumed by internet. The cost of energy to power data Centre gets doubled after every 5 years. The amount of power consumed by data Centre grew by 56% between 2005 and 2010. In 2010 approximately 1.1% to 1.5% of the total world energy was consumed by data Centre. In 2011 energy consumed by data Centre was approximately 1,00,00,000 MW, which has generated 40,568,000 tons of CO2 emissions. The data Centre consumes only 20-30 % to operate, whereas rest of the 70-80 % of the energy consumption is wasted due to over-provisioned idle resources which approximately results to 20,000,000 tons of CO2 emission. In Cloud Computing to store any document consumes less power whereas conventional computing consumes much more power as compared to it8. Electricity is needed to operate servers, interconnecting telecommunication networks and to cooling of system. Data Centre is not much costly to build but not eco-friendly for the environment9. Reducing energy consumption is a challenging issue. Even government is also pressurizing to reduce power consumption to reduce CO2 emission and greenhouse effect. Because of this reason Google, Yahoo is building their Data Centre in Barren desert surrounded by Columbia River in US in order to obtain cheap hydro power. From total expenditure spend to work efficiently Amazon EC2 pays 42% for energy usage. Data Centre named Microsoft Dublin consumes approximately 5.4 MW of energy. A lot of research has been done to gather different aspects for managing and consuming energy used in data centers.

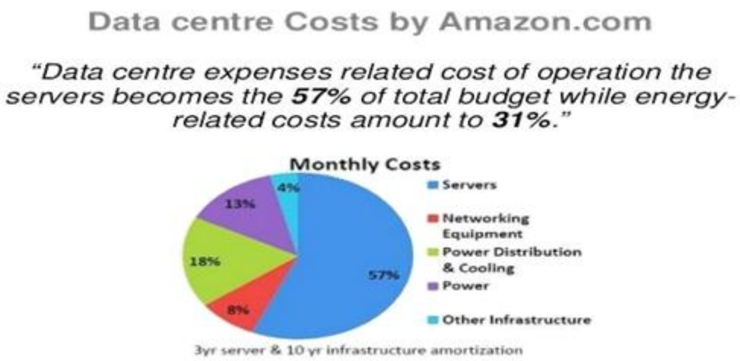


Figure1. Energy distribution in Amazon DATA center

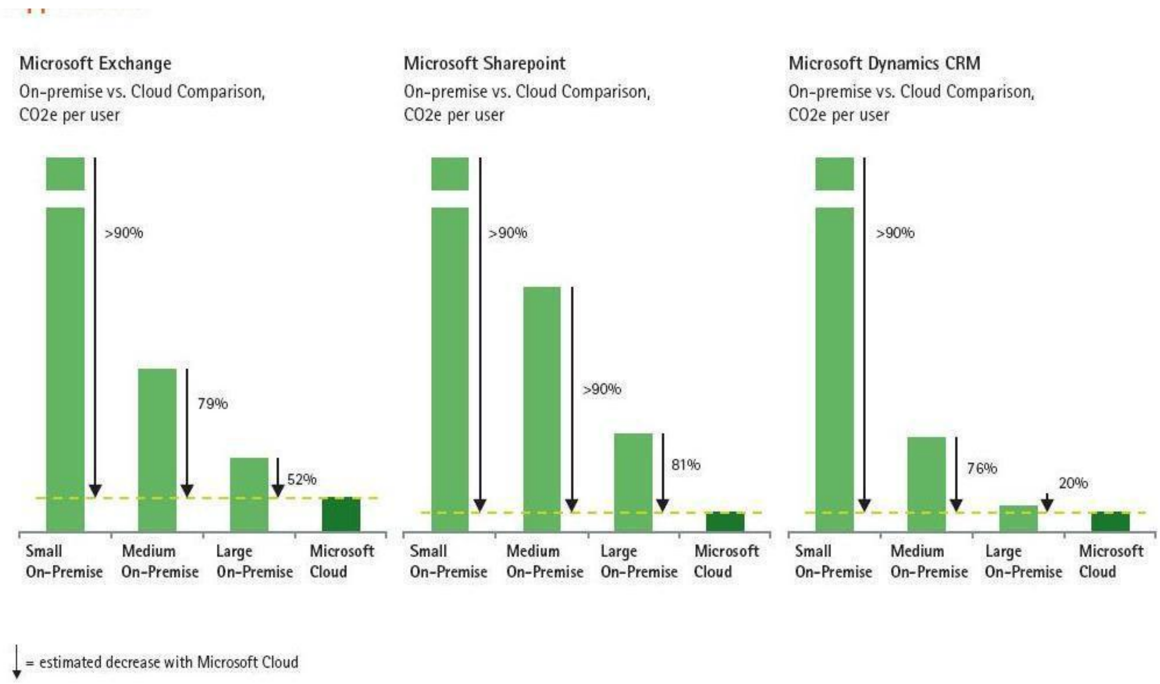


Figure 2 .Comparison of carbon Emissions of cloud-based vs. On Premise delivery Of Three Microsoft Application

Energy Efficiency in ICT (Information and communication technology) Infrastructure:-

ICT consumes an increasing amount of energy from productivity and economic prosperity and in reducing energy from other sources are through e-work and e-learning.

Network design is made to minimize the infrastructure costs and maximize the Quality of Service (QoS).As well as using of Smart devices in home and office's is also playing a complex route in energy consumption to optimize energy management .

Thus ICT can reduce energy consumption and carbon emission but this reduction is partially offset by the power used by data centers and computer network all over the worlds.

Energy-aware data centers:-

The key current technology for energy-efficient operation of servers in data centers is virtualization. VMs that encapsulate virtualized services can be moved, copied, created and deleted depending on management decisions. Consolidating hardware and reducing redundancy can achieve energy efficiency. Unused servers can be turned off (or hibernated) to save energy. Some hardware gets higher load, which reduces the number of physical servers needed. However, the degree of energy efficient self-management in data centres is still limited today.

Services should not only be virtualized and managed within a data center site but they should be moved to other sites if necessary. Not only the aspect of load has to be considered, also the 'heat' generated by a service has to be measured and accounted for before migrating operations. Every operational physical node produces heat. When a particular node is excessively used or is near other high-loaded nodes, hotspots can appear in a given data center. To avoid such hotspots, heat can be distributed across sites. Furthermore, services can be moved from sites with high load or high temperature to sites with smaller loads and lower temperatures. Generally, services should be moved to those locations, where they can operate in the most energy-efficient way. This kind of energy-efficient management of resources has to be realized by an autonomous energy management that is as transparent as possible to the user of a service.

- **Energy savings in networks and protocols :-**

Research has shown that communications, in particular, is one of the largest consumers of energy, however, energy optimization for communications must deal with the trade-offs between performance, energy savings and QoS . Some hardware already offers features that create an opportunity for energy-efficient operation such as turning off network interfaces and throttling of processors. Network protocols could also be optimized, or even be redeveloped in a way that enhances the energy-efficient operation of the network elements. Network devices could be enabled to delegate services to other devices so as to transfer services from energy inefficient to more energy-efficient devices or to devices that need to be always on, while certain other devices are turned off. The delegating device can then become dormant and be turned off. Currently, many basic network services have to remain active to periodically confirm their availability even when no communication is taking place. These 'soft states' make it impossible to turn off certain system components; therefore, new protocols need to be designed to work around such soft states so as to increase the energy efficiency of the network.

Signaling can also be revisited in this context; whereas data and signaling traffic vary widely, the same technology and mechanisms are used for both (in so-called in-band signaling).While signaling needs only low bandwidth but can occur anytime, data traffic occurs after signaling has taken place, usually requires high bandwidth and traverses all network layers up to the application layer, and uses processing power through multiple layers of the network. Therefore the use of out-of band signaling should also be evaluated to design and improve energy-aware communication protocols.

- **The effect of Internet applications: -**

Thus far, we have considered the opportunities offered by cloud computing as a possible foundation for energy-efficient ICT infrastructures but have not discussed the nature of the applications themselves. We note that one large application area for the Internet is in information dissemination. From digital cameras embedded in mobile phones to environmental sensors to Web 2.0, end users are generating and interconnecting unprecedented amounts of information and this trend is expected to continue unabated. However, the professional, expedited and reliable distribution of content requires increasing investments in infrastructure build-out and maintenance, and a matching electricity bill to run the underlying ICT. Web, peer-to-peer and web-based video-on-demand services currently dominate Internet traffic and, taken together, consistently comprise 85% or more of the Internet traffic mix for several years. In practice, dissemination networks operate using methods and paradigms based on remote-access, replicating functionality in several parts of the protocol stack, and fail to benefit from recent advances in wired and wireless communications, storage technologies and Moore's law. If cloud computing becomes a significant platform for producing and accessing information, the amount of data that will be transferred over the Internet will increase significantly. Content replication and dissemination algorithms will then need to consider energy as a key parameter of optimal operation, and therefore cloud computing calls for a thorough re-examination of the fundamentals of major computation/communication/storage and energy/performance trade-offs.

Overview of Cloud Computing:-

Cloud computing has become an important lesson in our coming New Generation because it offers dynamic, high-capacity computing capabilities, including access to data archiving and complex applications, without

Requiring any additional computing resources. In Cloud computing It uses cloud data centers through virtualization technologies to offer a powerful and adaptable computer environment. This concept is widely promoted and developed, and gained the interest of many organizations, mainly due to the reduction in

Expenses or cost with time consuming capability which could be achieved by diminishing the investment in hardware and software. Cloud computing is "an old idea whose time has (finally) come" (p. 2). Service-Oriented Architecture (SOA), Micro service Architecture, parallel computing, distributed computing and grid computing, virtualization, and containerization are the basic concepts of cloud computing. Some of them are older, such as parallel computing, distributed computing, and virtualization; others are newer, such as SOA, Micro service Architecture, grid computing, or containerization. Cloud computing solutions are extremely dynamic. According to **Heininger**, the following keywords characterize this new ICT provisioning model offered by cloud computing: ubiquitous, service-centric, scalable, consumption-based and self-service. The concept is defined mainly by its characteristics. The National Institute of Standards and Technology (NIST) has presented cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" (p. 2).

According to Buyya et al. (p. 3) "a cloud is a type of parallel and distributed system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resource(s) based on service-level agreements established through negotiation between the service provider and consumers". Cloud computing integrates existing technologies and models to optimize the use of physical and logical resources. The resources are treated as services and are available to users according to their requirements. There are three main models: IaaS (Infrastructure as a Service), PaaS (Platform as a Service), and SaaS (Software as a Service). IaaS and PaaS provide services to independent software vendors and developers, while SaaS provides services to end users. A typology of cloud computing should consider the degree of accessibility it offers so that it can be ranked as private, public, hybrid, and/or community.

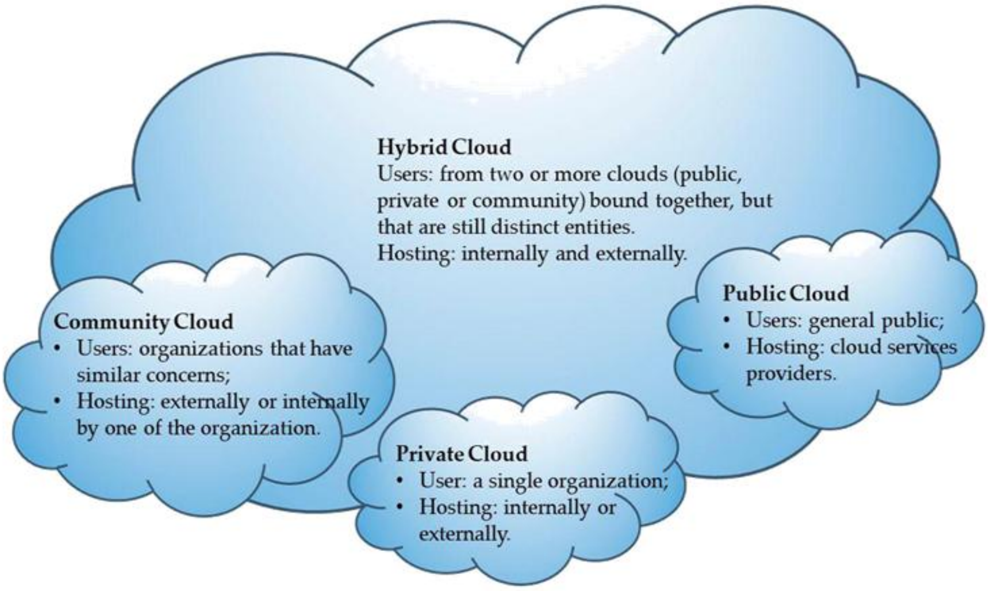


Figure. Types of Cloud Computing

According to Kliazovich et al. (p. 2) and with regard to the topic of this paper, "from the Energy efficiency perspective, a cloud computing data center can be defined as a pool of computing and Communication resources has organized in the way to transform the received power into computing or data Transfer work to satisfy user demands at demand of low cost".

This definition refers to the energy efficiency of the IaaS model. SaaS also provides benefits for environmental protection: through centralization of processing and Service sharing, it consolidates data center operations in order to use less equipment. SaaS providers could offer green software services deployed on green data centers with less replications or they could use algorithms that improve software energy efficiency without violating Service Level Agreements (SLAs). The cloud providers have more resources and more motivation than individual users have to invest in environmental protection. In the case of PaaS, the providers could offer facilities such as green schedule and green compilers. To help environmental protection through green cloud computing, both SaaS and PaaS providers have methods and tools to achieve software-level energy optimization. The increase in the popularity of cloud technology was due to the benefits it brought to individual consumers and companies. These benefits include: flexibility, disaster recovery, reduced investment in ICT resources, optimized collaboration between members of an organization, and automatic software updates. Cloud computing is attractive to business owners, due to the possibility of dynamically increasing the resources accessed to match increases in the company's activities. For the environment, the advantages of cloud computing are: better strategies for energy efficiency, and reduced equipment requirements and lower CO2 emissions, with, consequently, less e-waste.

In order to switch to cloud computing, enterprises might also face the challenges of a change of Software/hardware architecture, obstacles to data transfer, and concerns about interoperability.

These technologies carry some risks, mainly related to security issues. In spite of this, cloud computing technologies are constantly growing as a result of the major benefits they offer to companies, i.e., access to high-performance computing resources and high-capacity storage together with lower costs. With regard to the influence on the environment, the sections below present in detail the main problems **identified in both the academic and the non-academic studies.**

RELATED WORK

In spite of the universal research consideration committed to control proficient asset distribution in distributed computing frameworks, it comes up short on the ideal unique power the board down to earth stages. Dynamic power the executives procedure requires a gauge of the outstanding burden of distributed computer have contemplated stochastic demonstrating of distributed computing frameworks to foresee the accessible assets and the remaining burden of the DC. Be that as it may, either accurate examination hold the prohibitive circulations for example, Poisson and Exponential, are utilized for the entry what's more, flight rates of the cloud outstanding task at hand or the exactness of the examination is debased by certain approximations. Distinctive prescient arrangements endeavor to foresee the solicitation rate and to follow the future heaps of the DC investigate the viability of VM and host asset usage expectations in the VM combination task utilizing genuine remaining burden follows. The proposed methodology gives a considerable improvement over other heuristic calculations in diminishing vitality utilization, various VM movements and number of SLA infringement. Dynamic Virtual Machine (VM) solidification is a standout amongst the most encouraging answers for lessen vitality utilization and improve asset usage in server farms. Since VM union issue is carefully NP-hard, numerous heuristic calculations have been proposed to handle the issue. Present a half and half improvement demonstrate that permits a cloud specialist organization to build up virtual machine (VM) arrangement methodologies for its server farms so that vitality effectiveness and system nature of administration are mutually upgraded. For the most part, VM situation is a movement not completely incorporated with system activities. The proposed system permits cloud suppliers to achieve a harmony between the vitality productivity of their frameworks and the system nature of administration they offer to their clients. Maurizio Presented another system to diminish carbon dioxide discharges in combined Cloud environments. All the more explicitly, we propose an answer that enables suppliers to decide the best green goal where virtual machines ought to be moved so as to decrease the carbon dioxide outflows of the entire united condition. Discusses the multiplication of Cloud administrations have incredibly affected our general public, how green are these administrations are yet to be replied. Despite the fact that request acceleration for green administrations has become because of societal mindfulness, the ways to deal with give green administrations and build up Green SLAs stay negligent for cloud or framework suppliers. The fundamental test for the cloud supplier is to oversee Green SLAs with their clients while fulfilling their business targets, for example, augmenting benefits by bringing down consumption for efficient power vitality. Presented that with quickly expanding interest for distributed computing innovation, vitality proficiency has turned out to be profoundly imperative in distributed computing foundations. Distributed computing idea offers minimal effort and abnormal state of accessibility. Be that as it may, regardless it makes them challenge issues, for example, asset the board and power utilization. In this idea, diminishing vitality utilization and amplify asset use, turned into an essential worry of numerous asset the board techniques. Presented framework that handles ongoing and non-continuous assignments in a vitality proficient technique without bargaining much on neither unwavering quality nor execution. Of the three processors, two processors for example the first and second, handle ongoing undertakings, utilizing soonest Deadline-First (EDF) and Earliest-Deadline-Late (EDL) booking calculations individually. On the third processor, the non-continuous assignments are planned utilizing the First Come First Served (FCFS) booking calculation and the errands are kept running at a vitality effective recurrence. Our recreation results show critical vitality reserve funds contrasted with the current Stand-by Sparing for Periodic Tasks (SSPT) for a couple of execution situations. States that registering is a computational structure that gives a gathering of virtualized assets as Service. Distributed computing is exceedingly beneficial financially savvy benefits in the business world in the present day situation. Be that as it may, the vitality utilization of Data Centers is a major issue developing out of developing interest for cloud administrations. That sort of basic issues diminishes the net revenue, yet additionally the essential standards of Green figuring is coordinated towards condition amicable calculation. Discussed that distributed computing is another administration display for sharing a pool of processing assets that can be quickly gotten to and discharged dependent on a merged framework. Previously, singular use or organization can just utilize their very own servers to oversee application projects or store information. In this way it might cause the predicament of complex administration and weight in "possess and-use" designs. To fulfill unsure remaining burdens and to be profoundly accessible for clients anyplace whenever, giving more assets are required. Thus, asset over-provisioning and repetition are basic circumstances in a customary working framework.

Green Cloud Computing in Non-Academic Studies:-

In non-academic studies, both the negative and positive effects of cloud computing on the ecosystem have been studied and highlighted in various moments by different authors and organizations. According to Mines cloud computing infrastructure has two critical elements regarding its influence on the environment: energy efficiency and resource efficiency. According to the same author, cloud computing characteristics bring benefits to environmental protection even when they don’t have this scope explicitly. Virtualization and automation software increase energy efficiency and decrease the physical footprint and e-waste. The same activities will be performed with less equipment and less energy consumption both for users and green data centers, due to enhanced resource efficiency. Pay-per-use and self-service encourages users to limit consumption to real needs. Shared resource access allows users (organizations or individuals) to use the same infrastructure and services. This characteristic will decrease the amount of equipment required. Demand for cloud computing services is expected to continue to increase. This will increase the energy consumption and equipment, without the adoption of appropriate measures to protect the environment by cloud computing providers that will make the cloud greener. Figure 5 presents these influences. In this case, most of the studies concentrate on energy consumption in cloud computing. In a report from Accenture, it is claimed that, by energy optimization, CO2 emissions in cloud computing might be reduced by at least 30%. Another study, undertaken in 2011 by an independent company on several multinational companies that had been using cloud computing for at least two years, estimates that the energy saving for US companies would be more than $12 billion per year. The annual carbon reduction is estimated, in the same study, at 85.7 million metric tons—equivalent of 50% of CO2 emissions.

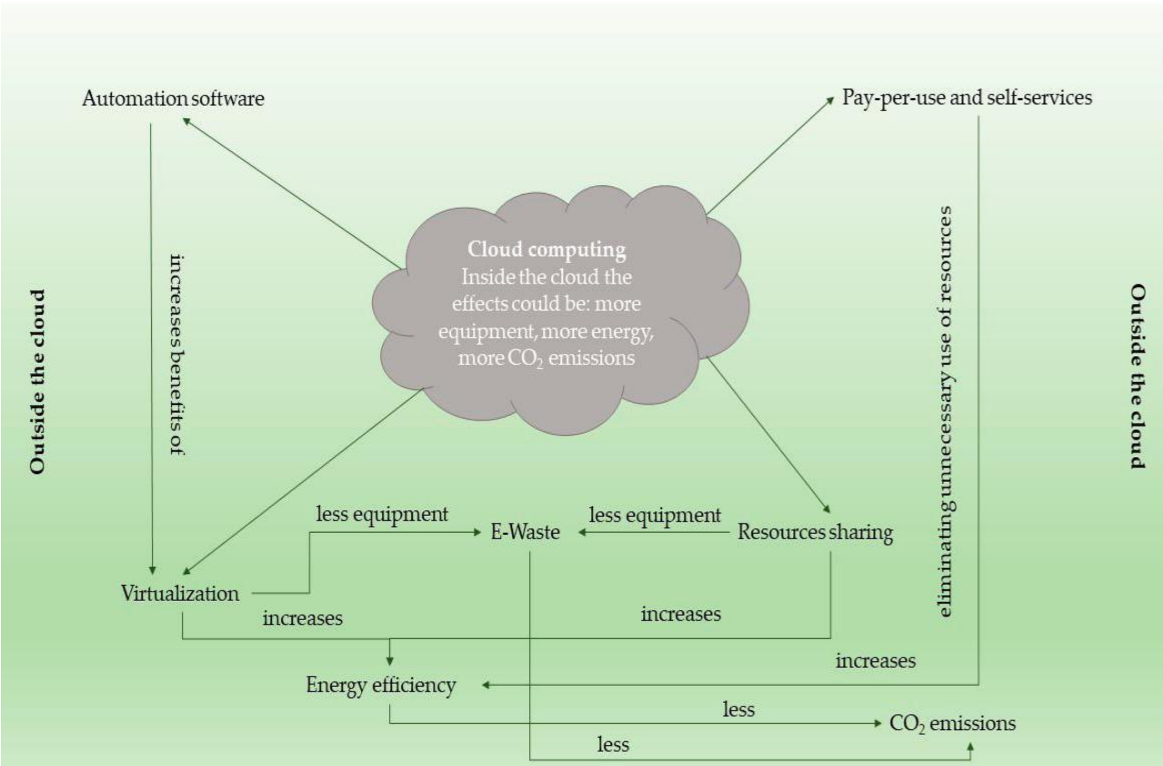


Figure 5. Cloud computing characteristics and their influences on the environment.

The cut in energy consumption and the required hardware is accompanied by reducedCO2 emissions, as well as reduced (e-) waste. An important role is played by the major cloud companies—Apple, Facebook, Google, Amazon, Microsoft, IBM, Salesforce, etc.—which have committed to use only renewable energy in data centers, and are showing early signs of accomplishing this promise .

VII. Scheduling

In Cloud Computing, scheduling directly affects some important parameters of cloud environment like energy efficiency. The major components of energy consumption Data Center are server, cooling system and interconnecting telecommunication. So to reduce energy consumption of Data Center is to decrease the number of server which is in active state to receive and process tasks10. This can be possible by effective scheduling which analyses both the load on network link as well as the occupancy of outing queue at network switches. Figure 4 Resource scheduling: The VMS are allocated to the task for the processing according to the CPU time or the capacity of the resource to execute the workload in given time period. To overcome the problem of delaying the task assigns b the broker

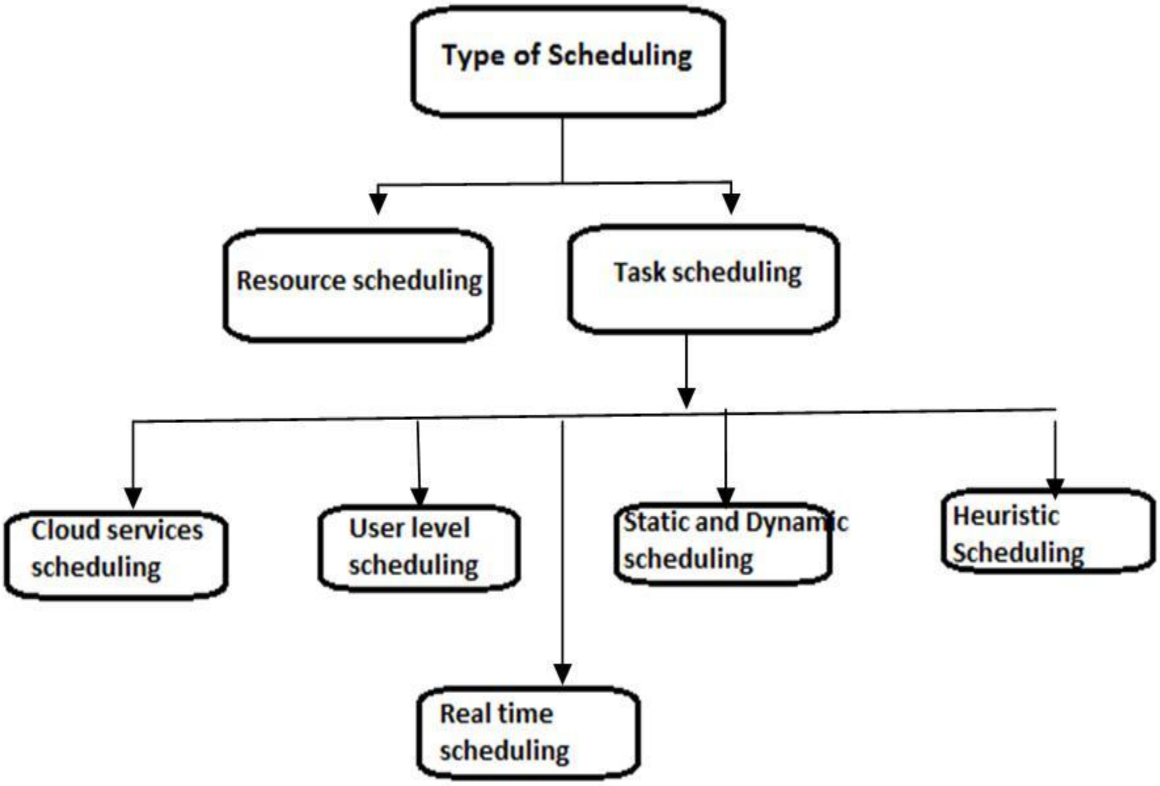


Figure Task Scheduling

Task Scheduling

In task scheduling some tasks are consolidated to form one job for further assignment to the processors. Task scheduling is of five types:

- **Cloud service scheduling:**

This scheduling is categorized at two levels: 1. User level scheduling: It deals with issues raised by the service management between service provider and customers. 2. System level scheduling: It manages resources with the Data Center.

- **User level scheduling:**

It deals with the dynamically fluctuating resources demand. In the cloud environment resources are of distinct type and the main focus of this scheduling is to give high SLA to the customer with the minimum cost.

• Static and dynamic scheduling:
In the static scheduling the information of the task arrived for the execution is well known beforehand. But in dynamic the knowledge about task is hidden. The dynamic scheduling is the one of big challenge to overcome.

- **Heuristic scheduling:**

This scheduling is the major concern to improve the energy efficiency in cloud data centers. This scheduling need algorithms which can schedule the task over VM on the basis of artificial intelligence or we can say the it assign task to the resources as like human being thing by checking all the factor to take care of all the parameters.

5. Task Consolidations

The workflow sim simulator is used to implement the technique in which clustering is applied on the cloudlets to collect same properties task on VM to decrease the data transmission over the different severs. The jobs with the particular configuration are arrived through the broker from the client to get processed. Then methodology will apply on the cloudlet to assign Virtual Machine for processing in minimum time slot. The scheduling techniques are implemented to utilize the resources effectively and the delay between task executions can remove completely. But sometime wrong strategy make the negative effect and increase the congestion of tasks which raise the energy consumption and put bad impact on other parameters also. The job comes with the input data file size and output data file size on the basis of which we measure the impact factor of the job. The impact factor is job run time, it is measured by the CPU itself automatically which is known as burst time of the cloudlet. The existing model is based upon the imbalanced metrics to optimize the task scheduling over the scientific workflow executions over the cloud environments. The existing model utilizes the task clustering mechanisms to cluster the small tasks into the task group appearing like the single task to process them smoothly and quickly.

Objective

This paper mainly tackle the problems in relation to an energy efficiency in cloud computing. Major challenges in this research are the development of a system which consumes less energy and less cost.

RESEARCH METHODOLOGY

Power sparing methods (dynamic recurrence scaling and dynamic voltage scaling) used to spare power in inserted frameworks including phones. This kind of intensity sparing is not quite the same as what the majority of us for the most part think Dynamic Voltage and Frequency Scaling (DVFS) depicts the utilization of two about like reserve or sleep control states. These are valuable obviously. You can lessen the power devoured by your inserted framework by bringing down the recurrence or potentially voltage of the CPU and connected peripherals.

1- All the assignments given by the client are gotten and are isolated into 3 zones relying on the length of their work.

2- Zone A: It incorporates every one of those errands that include ordinary handling

3- Zone B: This zone contains every one of those errands that include input/yield preparing and consequently require more CPU than zone A

4- Zone C: This zone contains the interactive media employments and these errands require extremely high CPU than some other classes.

5- All the virtual machines or we can say the assets will work in 3 modes:

6- Low Mode/Basic Mode: The virtual machine will work in low-recurrence mode and a lesser measure of voltage will be provided

7- Medium Mode: Medium recurrence and medium voltage is provided

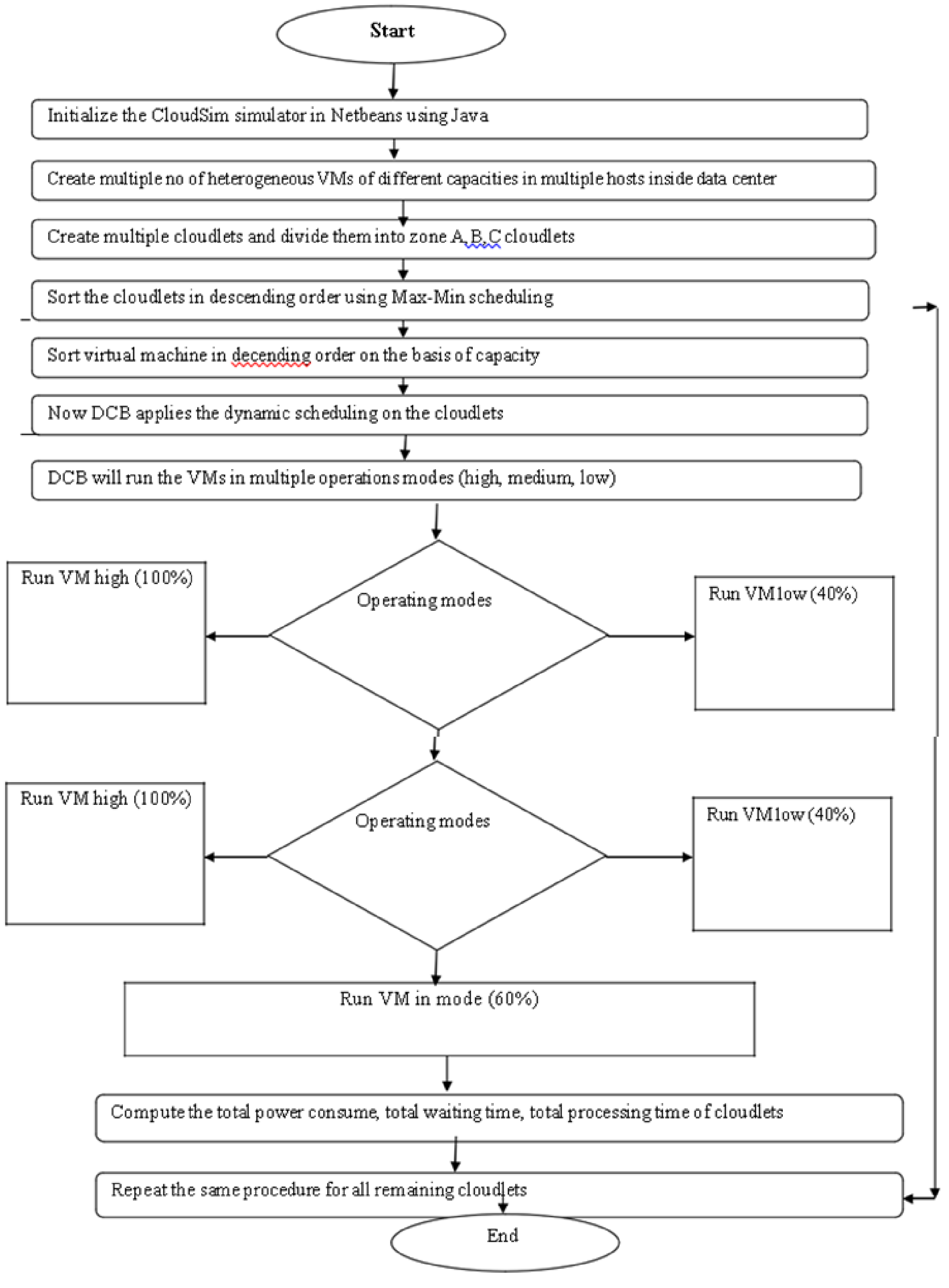
8- High Mode: High recurrence and higher measure of voltage are provided.

9- Assign every one of the errands in zone A to the machines that are working in the essential mode or low mode.

10- Assign every one of the errands in Zone B to the machines that are working in the medium mode.

11- The errands of Zone C are doled out to the machines that are working in high mode

(100% CPU use.



NetBeans:
NetBeans is an integrated development environment for developing primarily with Java, but also with other languages. He NetBeans IDE is written in Java and can run on Windows, OS X, and Linux. He NetBeans Platform allows applications to be developed from a set of modular software components called modules. Applications Based on the NetBeans Platform including the Netbeans IDE itself can be extended by third party developers.

Implementation of the Proposed Technique

The technique has been implemented the heterogeneous workload in CloudSim Toolkit by using Netbeans We have taken a different number of processors, different number of tasks and execution times for each task in each processor as input. Thesis shown in figure

Experimental result

Precisely Objective is to achieve accuracy, the performance of servers and the cloud environment can be maintained. Various experiments have been conducted and the results of existing work and the proposed work have been mentioned in Table 1 and Table 2 and Table 3

Here used three parameters used-

1-Total processing time

2- Total waiting time

3- Total processing cost

(I).**Total processing time**

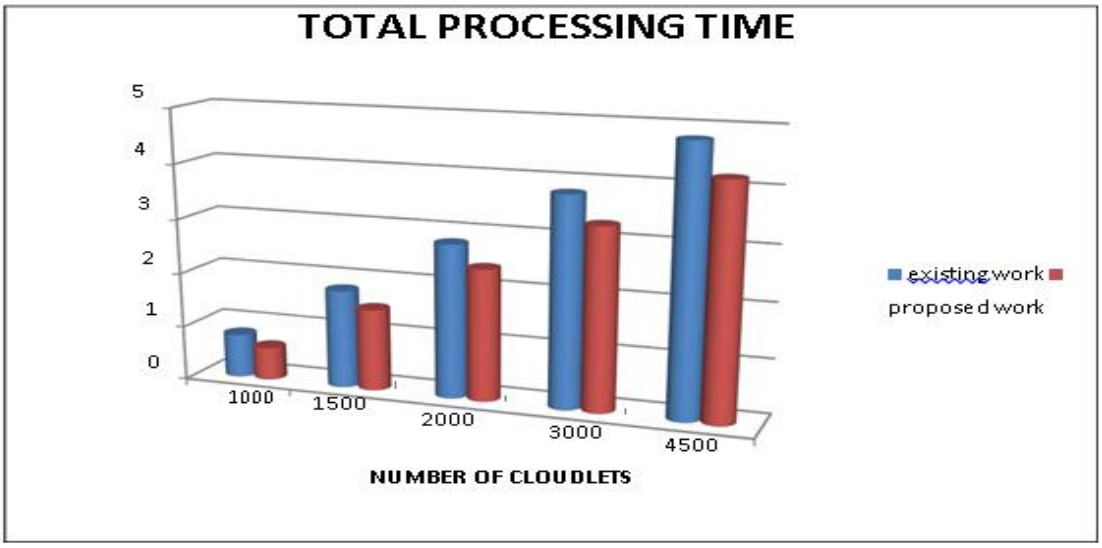
It is defined as the time interval between the request sent and response received by the cloud user/consumer.

Overall processing time is calculated as given as follow:-

Processing Time = Finish time of execution - Start time of execution PT=FT-ST

Table 1. Base vs Proposed Total Processing Time

S.NO	NO OF CLOUDLETS	TOTAL PROCESSING TIME OF EXISTING WORK	TOTAL PROCESSING TIME OF PROPOSED WORK
1	1000	1498667	1160711
2	1500	2250000	1743199
3	2000	2998667	2322844
4	3000	4500000	3486399
5	4500	6750000	5229598



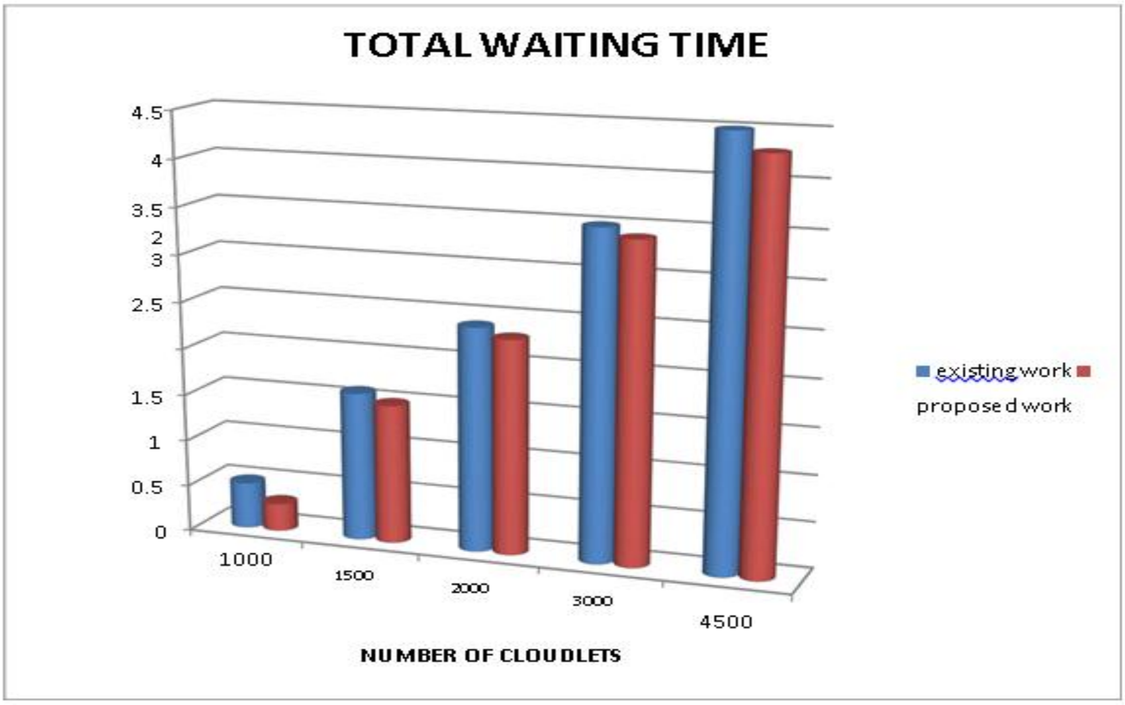
From the above bar chart, it is clear that the total processing time has been reduced.

(II) TOTAL WAITING TIME

Waiting Time = Allocation Time – Generation Time

Table 2. Base VS Proposed Total Waiting Time

S.NO	NUMBEROF CLOUDLETS	TOTAL-WAITING TIME OF EXISTING WORK	TOTAL-WAITING TIME OF PROPOSED WORK
1	1000	149056667	140588186
2	1500	336083333	317101126
3	2000	598111333	563857671
4	3000	1347166667	1270147745
5	4500	3033250000	2859139766

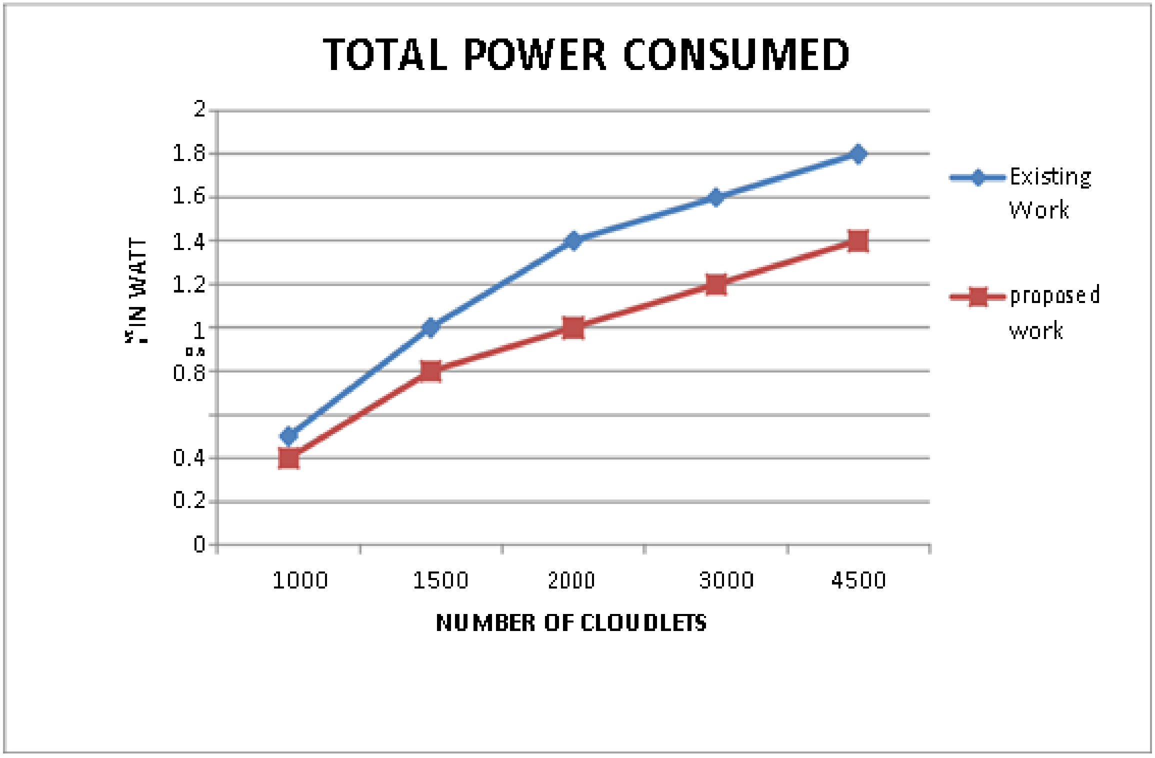


(III) TOTAL POWER CONSUMED

Total Power Consumed = Total Processing Time * Power of Virtual Machine

Table 3. Base vs. Proposed Total Power Consumed

S.NO	NO OF CLOUDLETS	TOTAL POWER CONUMED OF EXISTING WORK	TOTAL POWER CONSUMED OF PROPOSED WORK
1	1000	332704000	148571008
2	1500	499500000	223129469
3	2000	665704000	297323976
4	3000	999000000	446259057
5	4500	1498500000	669388528



From Application Attributes and Green Business Patterns to Energy Efficient Cloud Patterns

In the previous sections, the concept has been explained in two steps. To connect all information mentioned above and give a clue to what the concept (system) does, a use case scenario, a process flow diagram, and conceptual module have been added:

- Use Case Scenario

1- A user selects some application attributes and prioritizes them.
2- The cloud migration decider applies the migration algorithm and returns the migration decision.
3- If the migration is not recommended, user can reselect the attributes.
4- If not, user selects the green business process patterns that user wants to apply to the process or activity as illustrated in figure.

Figure: Green Process Patterns Selection

5-A set of correlated cloud patterns to the selected green patterns are selected based on the correlation matrix Table 3. The set can also be enhanced to match the user's requirements (attributes) by using the correlation matrix Table.

6-Cloud storage and communication patterns can be returned to the user in addition to the architectural patterns resulted from the previous step. This operation depends on the correlation matrix between different cloud patterns Table.

- **System Process Flow Diagram**

The following is the process flow diagram of the recommender systems. The diagram shows the border between the two steps explained in the concept as well as the sequence of processes done.

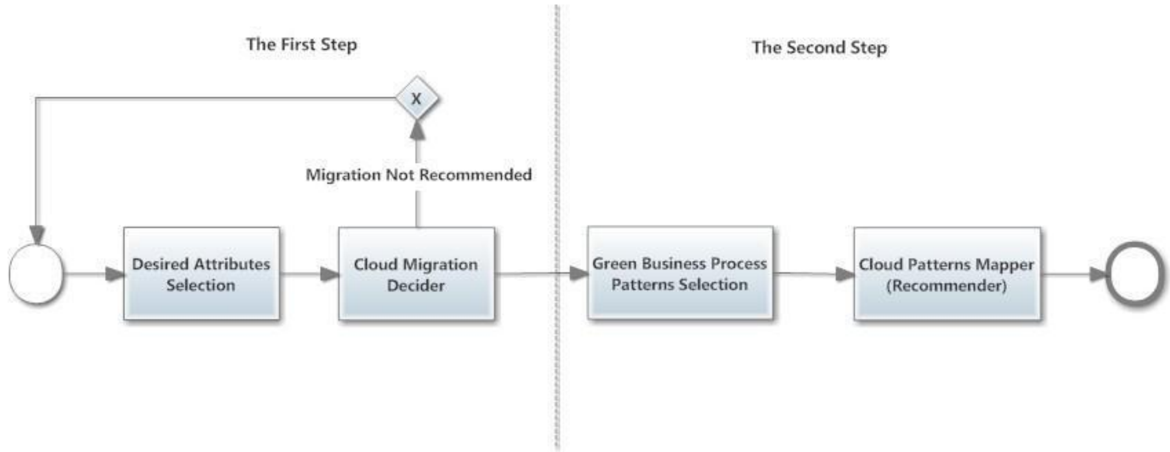


Figure: Process Flow Diagram of the Recommender System

System Conceptual Model

The following diagram shows the relations between applications attributes, cloud application patterns, cloud services, and green business process patterns.

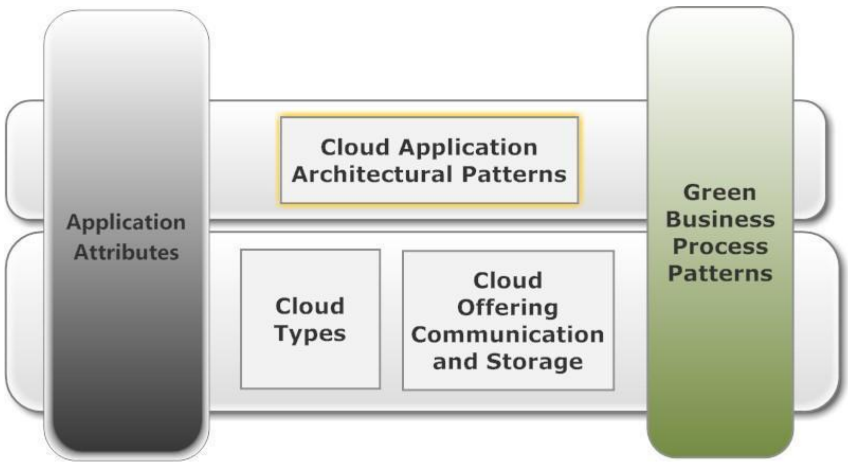


Figure 3.6: Conceptual model

The concept introduced in the chapter 3 is translated into a web application recommender system. The recommender system has been developed takes the selected application's attribute from the user as inputs, and returns to the user the recommended energy-efficient application architectures, storage patterns, and cloud types. The same as most of the web-based application, the recommender system is a three-tier application: user interface (UI), server, and database. In this chapter we will introduce in details the architecture of the recommender system, development environment, technology used, and some diagrams.

Client Side

As known in web applications, a web client application typically has the user interface (UI) of the application and has some form of programmatic control. It runs within the browser, and is responsible to interact with the user actions as shown in the sequence diagram Figure.

Through the Web UI provided in the recommender system, user can easily select the desired application attributes and prioritize them as shown in Figure 3.2 as well as the desired green pattern as illustrated in Figure

3.4. Some UI components have been provided in the Web interface to show the results. In this section, we will introduce the client part of the recommender system such as technologies used and other technical issues.

Technology Used

The user interface in the recommender system has been developed using Flash technology powered by Adobe Flex platform. The reasons behind this selection are several; the main reason is that this technology enables developer to build easily various Internet Applications using very flexible object-oriented scripting language Action Script 3. With these technologies, it is feasible to build a dynamic user interface, e.g., number of application's attributes can increase or decrease which leads to continuous changes on the UI, and so we need flexible technologies to perform this functionality and make any further enhancement to the system UI not very complex. We will add some information about the technologies that we used in the client side.
Flex Technology: We have developed the client side (Web User Interface) using Flex 4.6 SDK. Adobe defines flex as follows [61]: "Flex is a powerful open source application framework that allows you to easily build mobile applications for iOS, Android, and BlackBerry Tablet OS devices, as well as traditional applications for browser and desktop using the same programming model, tool, and codebase". Flex SDK is the software development kit for development and deployment of Rich Internet Application based on the Adobe Flash platform.

Flex is commonly compared to some technologies such as Ajax, XUL, JavaFX, and HTML5. The typical workflow for Flex is as follows:

- Define an application interface using a set of pre-defined components (form, buttons etc.)
- Build a user interface design by arranging the pre-defined component.
- Use the styles to define the visual design.
- Add the dynamic behaviour by a scripting language.
- Define and connect to data services or web services as needed.

To write applications in Flex, Flex framework provides two languages: MXML and Action Script. MXML is an XML-based declarative language used to layout user interface (UI) components. MXML can be used in non-visual aspects of an application, such as defining data bindings between user interface components and data sources on the server. Action Script is used as a scripting language with MXML. It is an object-oriented language and a dialect of ECMAScript (a superset of syntax and semantics of the languages as JavaScript). The compiled file from the source code contains mxml and action script is a SWF file (Shockwave Flash) that can be embedded inside normal HTML pages or JSPs.

Technical Issues

As shown in the sequence diagram Figure 4.3, when user selects desired attributes; the "programmatic control" that is written in ActionScript takes these selected attribute, stores them, and prepare the parameters to send to the server. The parameters include the application attributes and their priorities (low, high, and critical) are sent to the server as pairs from the attribute and its priority. To connect the client side with the server, HTTPService connection component has been used. HTTPService is a defined component in ActionScript 3 used to connect to any severs technology. It handles the request object that will be sent to the server as well as the response object that will be received from the server. HTTPService is defined in Flex code as follows:

```
<s:HTTPService  id="httpXmlDataService"
                url="http://localhost:8080/ThesisWork/CloudDecisionRecommender" resultFormat="e4x"
                result="resultHandler(event)" fault="faultHandler(event)"
                useProxy="false" method="POST" />
```

The "url" attribute represents the address where the request will be sent to be processed. In the code snippet shown above, the "url" attribute denotes to Cloud Decision Recommender which is the program that will execute the migration algorithm as shown in the sequence diagram Figure 4.3. The "result" attribute has the name of the function that will handle the response when comes from the server. It is also shown in the code snippet that POST method has been used to fix the issue of query string length when GET method is used. The response return in an XML data format; and the data is parsed in the function defined by "result" attribute. The same sequence is followed when the user selects the desired Green Process patterns to get the recommended related Architectural patterns.

Server Side

In the server side, the business logic of the recommender system has been implemented. The business logic (1) exchanges data between the database and the user interface, (2) implements the migration decision algorithm in section 3.1, and (3) retrieves the related patterns from the database, analyzing them, and returns the most relevant cloud patterns back to the client.

Technology used

Java Servlet programming language has been used in order to develop the recommender system server components. Servlet is a Java code that runs in a server application to answer client requests. Servlet uses the standardized JAVA APIs. Additionally, it has its own APIs such as HTTPServlet that used to read client requests and generate the response. Eclipse integrated development environment (IDE) has been used as a development tool for Java Servlet. Eclipse is an open source java application platform has major advantages over other IDEs. It is widely used among most companies nowadays.

Apache Tomcat has been used as a web server where the recommender system server components reside and run. , Apache Tomcat is an open source software implementation of the Java Servlet and Java Server Pages technologies. It powers numerous large-scale, mission-critical web applications across and diverse range of organizations.

Technical Issues

Two Servlet programs have been developed to execute the business logic of the recommender system: The first Servlet serves the first step of the concept that explained in section 3.1 which recommends the user with the migration decision. As shown in the sequence diagram Figure Cloud Decision Recommender Servlet executes the migration algorithm: it reads the client request, parses the parameters which include the attributes and their priorities, builds the queries that will be executed in the database, and finally it uses the weights and level values retrieved from the database to calculate the application score on which the migration decision depends. The second Servlet implements the second step which recommends the user with the energy efficient cloud patterns. As shown in the sequence diagram, Architecture Recommender Servlet reads the client request, parses the parameters which include the desired Green Business Process patterns, uses the selected attributes from the first Servlet, and builds the query that will retrieve cloud patterns that are relevant to both of green patterns and application's attribute.

Database

The place where all data includes the attributes, cloud patterns, business process patterns, and their relations used in the recommender system are stored. We preferred to save all the data required in the database even those used in the first step calculation to add more flexibility and extensibility to the system.

Technology used

MySQL Database has been used in the system. MySQL has many advantages mentioned in [63]: it is an ACID compliant database with full commit, rollback, crash recovery, and row level locking capabilities. MySQL also improves the scalability, performance, and delivers the ease of use that has made MySQL a preferable database solution. WampServer has been used as a database server. WampServer is a Windows web development environment. It allows the user to create MySQL database and to use PHP My Admin and SQLite Manage to easily manage the database.

The Entity Relationship Diagram

Figure is the entity relationships (ER) diagram of the database of the recommender system. It explains the relations between different kinds of patterns used in the system.

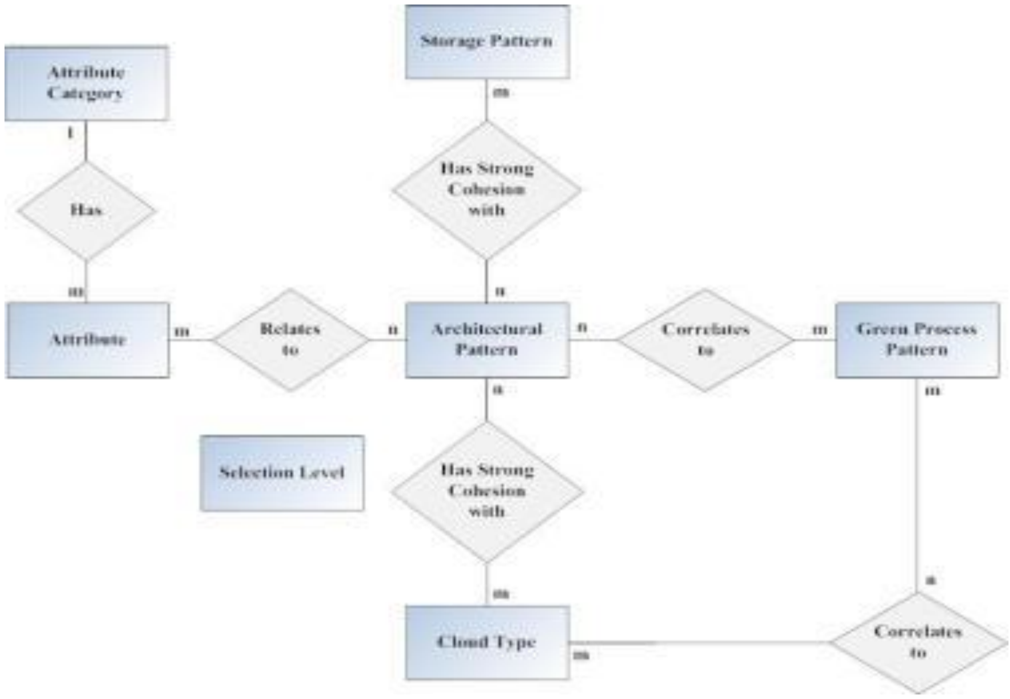


Figure: ER Diagram of the Recommender System Database

The Tables and the Attributes Description

In the previous subsection, the entity relationship diagram of the system's database has been presented. In this subsection, the description of the database's tables and attributes will be provided. The relations between some tables will be explained as they represent the relation tables shown in Chapter 3. These relations are which the concept primarily depends on. The following is the description of some of recommender system tables:

	Table	Table/Attribute Description
	ATTR_CATEGORY	Stores the high level application quality attributes information such as Data, Security etc.
Attribute	CAT_ID	Unique incremental number (Primary Key)
	CAT_NAME	String data represents the name of the attribute
	ATTRIBUTE	Stores the low level application's attribute for every high quality attribute.
	ATTR_ID	Unique incremental number (Primary Key)
	ATTR_NAME	String data represents the name of the attribute (Display Name)

	ATTR_NAME_ID	Summarized name used as an identifier (parameter names) in the query string that comes from the client request.
	DESCRIPTION	Textual data describes the application's attributes
	CAT_ID	Foreign key to join the attribute with category table
	WEIGHT	Signed value measures the contribution of an attribute in the Migration decision.
CLOUD_ARCHITECTURE_PATTERN		Stores the cloud architectural patterns data
	ID	Unique incremental number (Primary Key)
	NAME	Textual data represents the display name of the pattern
	DESCRIPTION	Textual data describes the pattern
CLOUD_STORAGE_PATTERN		Stores the cloud storage patterns data such as NoSQL and Blob storage
	ID	Stores the cloud storage patterns information

	NAME	String data represents the display name of the pattern
	DESCRIPTIO N	Textual data describes the pattern
	GREEN_PATTERN	Stores the green business process patterns data
	ID	Unique incremental number (Primary Key)
	NAME	String data represents the display name of the green pattern
	NAME_ID	Summarized name for the pattern used as identifier (parameter name) in the client request
	DESCRIPTIO N	Textual data describes the pattern
	CLASS	String data describes the class of the green pattern: basic pattern or process centric
	GREEN_ARCH_PATTERN	Junction table used to join the architectures with the green patterns tables such that translate the relations between patterns described in Chapter 3 to database records. Figure 4.2 illustrates one of the junction tables In the recommender system database.
	ID	Primary key of the record (Relation ID)
	GREEN_ID	Foreign key to join the green patterns table
	ARCH_ID	Foreign key to join the architecture table

	SELECTOR_OPTION	Stores the priorities of the application's attributes (Low, High, and Critical) and their numerical values that used in the migration Algorithm.
	ID	Unique incremental number (Primary Key)
	NAME	The name of the level (Low, High, Critical)
	VALUE	A numerical value for the level.

Table : Database tables and attributes Description

VIII. Conclusion

Green distributed computing is the system of actualizing the strategies to improve the proficiency of figuring assets and diminish the carbon impressions discharged by the IT server farms. Cloud Infrastructure is the most significant part in a cloud. It might comprises a huge number of servers, arrange plates and gadgets, and normally serve a great many clients internationally. Such an extensive scale server farm will devour tremendous measure of vitality. In this paper, we have proposed at ask effective power sparing calculation in cloud condition. We will actualizing the proposed system in the CloudSim test system by making diverse number of virtual machines and cloudlets. It will diminish the general vitality utilization and cost of utilizing the cloud. Power and vitality utilization are key worries for server farms. These focuses incorporate a great many server and bolster foundations for cooling also.. The proposed work hence advances a proficient vitality utilization strategy, remembering the issues defined in the current framework. In proposed system cloud condition is created in java, conveyed on CloudSim toolbox and the investigation results have been gone along according to quantitative examination. In proposed procedure, control sparing in green cloud condition has been finished utilizing Max-Min joined with DVFS (Dynamic Voltage and Frequency Scaling) so that the cloudlet can be isolated according to their need and distributed to processor to keep running at either at full recurrence or DVFS. The proposed proposition reviews Cloud Computing and its experience depicting the engineering, models and advantages moving further towards the idea of Green Cloud Computing as the vitality proficiency is one of serious issue with cloud computing. The proposed system indicates 55.3% decrease in power utilization, 5.7% decrease in absolute holding up time and 22.5 % decrease in complete preparing time. The advantage of diminishing force utilization is less warmth is created by gadget which benefits the mechanical plan just as it improves the lifetime time of gadgets.

Reference

[1]. B. Wadhwa and A. Verma, "Carbon Efficient VM Placement and Migration Technique for Green Federated Cloud Datacenters," IEEE, pp. 2297-2302, 2014.

[2]. C. Qiu, H. Shen and L. Chen, "Towards Green Cloud Computing: Demand Allocation and Pricing Policies for Cloud Service Brokerage," IEEE, pp. 203-2012, 2015.

[3]. Eduard OróVictorDepoorter,AlbertGarcia and JaumeSalom,"Energy efficiency and renewable energy integrationin data centres. Strategies and modeling review,"ELSEVIER,pp.1-17,2015.

[4]. F. Farahnakian, T. Pahikkala, P. Liljeberg, J. Plosila and H. Tenhunen, "Utilization Prediction Aware VM Consolidation Approach for Green Cloud Computing," IEEE, pp. 381-388, 2015.

[5]. L. A. Rocha and E. Cardozo, "A Hybrid Optimization Model for Green Cloud Computing," IEEE, pp. 11-20, 2014.

[6]. M. Giacobbe, A. Celesti, , M. Fazio, M. Villari and A. Puliafito, "An Approach to Reduce Carbon Dioxide Emissions Through Virtual Machine Migrations in a Sustainable Cloud Federation," IFIP, 2015.

[7]. M. S. Hasan, Y. Kouki, T. Ledoux and J. L. Pazat, "Exploiting Renewable sources: when Green SLA becomes a possible reality in Cloud computing," IEEE, pp. 1-14, 2015.

[8]. Marcus Hähnel, Björn Döbel, and Hermann Härtig. eBond: Energy Saving in Heterogeneous R.A.I.N. In e-Energy Conference.ACM, 2013.

[9]. J. Hamilton. Cooperative Expendable Micro-Slice Servers (CEMS): Low Cost, Low Power Servers for Internet-Scale Services. In Conference on Innovative Data Systems Research. Citeseer, 2009.

[10]. M.R. Hines and K. Gopalan. Post-copy Based Live Virtual Machine Migration Using Adaptive Pre-Paging and Dynamic Self-Ballooning. In International Conference on Virtual Execution Environments, pages 51–60. ACM, 2009.

[11]. T. Hirofuchi, H. Nakada, S. Itoh, and S. Sekiguchi. Enabling Instantaneous Relocation of Virtual Machines with a Lightweight VMM Extension. In *2010 10*th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, pages 73–83. IEEE, 2010.

[12]. Urs Hölzle. Brawny cores still beat wimpy cores, most of the time.IEEE Micro, 30(4), 2010.

[13]. Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjunsingh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, and Min Zhu. B4: Experience with a Globally-Deployed Software Defined WAN. In Proceedings of the ACM SIGCOMM *2013*

[14]. H. Jin, L. Deng, S. Wu, X. Shi, and X. Pan. Live Virtual Machine Migration with Adaptive Memory Compression. In International Conference on Cluster Computing and Workshops, pages 1–10. IEEE, 2009.

[15]. Thomas Knauth and Christof Fetzer. Energy-aware Scheduling for Infrastructure Clouds. In International Conference on Cloud Computing and Science. IEEE, 2012a. URLhttp://se.inf.tudresden.de/pubs/papers/knauth2012scheduling.pdf.

[16]. Brochure of VMWare Infrastructure 3, VMWare. http://www.vmware.com/pdf/vi_brochure.pdf. (Last accessed August 12,2009).

[17]. Economou, D., Rivoire, S., Kozyrakis, C. and Ranganathan, P.(2006) Full-System Power Analysis and Modeling for Server Environments. Workshop on Modeling, Benchmarking, and Simulation (MoBS), June, Boston, USA.

[18]. Chase, J.S., Anderson, D.C., Thakar, P.N., Vahdat, A.M. and Doyle, R.P. (2001) Managing Energy and Server Resources in Hosting Centres. Proc. 18th ACM Symp. Operating System Principles, Banff, Canada, October, pp. 103–116. ACM.

[19]. Chen, G., He,W., Liu, J., Nath, S., Rigas, L., Xiao, L. and Zhao, F.(2008) Energy-Aware Server Provisioning and Load Dispatching for Connection-Intensive Internet Services. NSDI'08: Proc. 5th USENIX Symp. Networked Systems Design and Implementation, San Francisco, CA, USA, April, pp.

[20]. Urgaonkar, B., Shenoy, P., Chandra, A., Goyal, P. and Wood, T. (2008) Agile dynamic provisioning of multi-tier Internet applications. ACM Trans. Auton. Adap. Syst., [40] Gelenbe, E. (2009) Steps towards self-aware networks. Commun.ACM, **52**,

[21]. Sarokin, D. (2007) Question: energy use of Internet.http://uclue.com/?xq=724. (Last accessed August 12, 2009).

[22]. Jia, X., Li, D. and Du, D. (2004) QoS Topology Control in Ad hoc Wireless Networks.Proc. IEEE INFOCOM'04, HongKong, China, March, IEEE.

[23]. M. A. Sharkh, M. Jammal, A. Shami, A. Ouda, "Resource allocation in a network-based cloud computing environment: design challenges," IEEE Communications Magazine, vol. 51, no. 11, pp. 46-52, November 2013.

[24]. Leon, L. Navarro, "Limits of energy saving for the allocation of data center resources to networked applications," IEEE INFOCOM, pp. 216-220, 10-15 April 2011.

[25]. B. Guenter, N. Jain, C. Williams, "Managing cost, performance, and reliability tradeoffs for energy-aware server provisioning," IEEE INFOCOM, pp. 1332-1340, 10-15 April 2011.

[26]. J. Doyle, R. Shorten, O'Mahony, D., "Stratus: Load Balancing the Cloud for Carbon Emissions Control," Cloud Computing, IEEE Transactions on , vol.1, no.1, pp.1,1, Jan.-June 2013.

[27]. J. Doyle, R. Shorten, D. O'Mahony, "Stratus: Load Balancing the Cloud for Carbon Emissions Control," IEEE Transactions on Cloud Computing, vol. 1, no. 1, pp. 1-1, June 2013.

[28]. Li Hongyou, Wang Jiangyong, Peng Jian, Wang Junfeng, Liu Tang, "Energy-aware scheduling scheme using workload-aware consolidation technique in cloud data centres," China Communications, vol. 10, no. 12, pp. 114-124, Dec. 2013.

[29]. L. Benini, A. Bogliolo, and G. De Micheli, "A survey of design techniques for system-level dynamic power management," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 8, no. 3, pp. 299 – 316, June 2000.

[30]. Y. Audzevich, P. Watts, S. Kilmurray, and A. W. Moore, "Efficient photonic coding: a considered revision," SIGCOMM workshop on Green networking (GreenNets '11), ACM, New York, NY, USA, 2011.

[31]. Bin Lin, Shanshan Li, Xiangke Liao, Qingbo Wu, and Shazhou Yang, "eStor: Energy efficient and resilient data center storage," 2011 International Conference on Cloud and Service Computing (CSC), pp. 366-371, December 2011.

[32]. Ruay-Shiung Chang, Hui-Ping Chang, and Yun-Ting Wang, "A dynamic weighted data replication strategy in data grids," IEEE/ACS International Conference on Computer Systems and Applications, pp. 414-421, April 2008.

[33]. Xiaowen Dong, T. El-Gorashi, and J. M. H. Elmirghani, "Green IP Over WDM Networks With Data Centers," Journal of Lightwave Technology, vol. 29, no. 12, pp. 1861-1880, June 2011.

[34]. Fan Ping, Xiaohu Li, C. McConnell, R. Vabbalareddy, Jeong-Hyon Hwang, "Towards Optimal Data Replication Across Data Centers," International Conference on Distributed Computing Systems Workshops (ICDCSW), pp. 66-71, June 2011.

[35]. Wenhao Li, Yun Yang, and Dong Yuan, "A Novel Cost-Effective Dynamic Data Replication Strategy for Reliability in Cloud Data Centres," International Conference on Dependable, Autonomic and Secure Computing (DASC), pp. 496-502, December 2011.

[36]. D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, A.Y. Zomaya, "Energy-Efficient Data Replication in Cloud Computing Datacenters," Springer Cluster Computing, 2014.

[37]. Cisco, 2011, *Global cloud index: Forecast and methodology, 2011–2016* (Whitepaper). Cisco.

[38]. D. G. d. Lago, E. R. M. Madeira, and L. F. Bittencourt, "Power-aware virtual machine scheduling on clouds using active cooling control and dvfs," in Proceedings of the 9th International Workshop on Middleware for Grids, Clouds and e-Science, ser.

[39]. MGC'11. New York, NY, USA: ACM, 2011, pp. 2:1–2:6.

[40]. S. Srikantaiah, A. Kansal, and F. Zhao, "Energy aware consolidation for cloud computing," in Proceedings of the 2008 conference on Power aware computing andsystems, ser. HotPower'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 10– 10.

[41]. J. Kim, M. Ruggiero, D. Atienza, and M. Lederberger, "Correlation-aware virtual machine allocation for energy-efficient datacenters," in Design, Automation Test in Europe Conference Exhibition (DATE), 2013, 2013, pp. 1345–1350.

[42]. Wang, Y. Yao, X. Wang, K. Lu, and Q. Cao, "Carpo: Correlation-aware poweroptimization in data center networks," in INFOCOM, 2012 Proceedings IEEE, 2012, pp. 1125–1133.

[43]. Wang, Y. Yao, X. Wang, K. Lu, and Q. Cao, "Carpo: Correlation-aware poweroptimization in data center networks," in INFOCOM, 2012 Proceedings IEEE, 2012, pp. 1125–1133.