Virtualization Approaches in Cloud Computing

K C Gouda[#1], Anurag Patro[*2], Dines Dwivedi [*3], Nagaraj Bhat[#4]

[#]*CSIR Centre for Mathematical Modeling and Computer Simulation (C-MMACS)*
Wind Tunnel Road, Bangalore-37, India
[*]*Department of Computer Science and Engineering, National Institute of Technology*
Rourkela, Odisha, India, 769008

***Abstract*— In the growing age of cloud computing demands in almost all the sectors like TV channel to online shopping to educational institution etc., the cloud infrastructure cost goes very high. Therefore, the concept of virtualization needs to be understand and implement in the cloud computing syterms, which enables the user as well as the owners for the better and robust management and usage of the cloud. In this paper the virtualization approach in the cloud computing environment are well presented with the concept of the cloud service models. The virtualization process and implementation with their advantages are documented and the different types of virtualizations are presented with some survey on the Indian data centre and the virtualization architecture. This approach needs to be evaluated in different cloud platforms for finding the cost effectiveness in the cloud infrastructure development.**

***Keywords*— Cloud computing, hypervisor, virtualization, data centres**

I. INTRODUCTION

Cloud computing is defined by the concept of web centred computers, services and resources that system developers use to implement compound web based systems. It basically deals in allocating the computing resources over internet .In cloud computing one uses the service over internet rather than keeping data in hard drive and regularly updating the applications. Thus it is observed that there is a significant change in workload. The computers make up the cloud handles the workload instead. The hardware and software demands are decreased. An interface software like web browser is the only thing required by the user's computers and rest care is taken by the cloud's network. Thus in easier words it can be called as "internet based computing".

I. CLOUD SERVICE MODELS

A cloud can interact with a client in variety of ways, through capabilities called services. There three major types of service, and the details are presented schematically in figure 1. The details discussion of each servise are discussed below.

A. SaaS(software as a service)

The software as a service (SaaS) provides cloud based foundation for software on demand. In general SaaS is a web delivered content that users access by a web browser. It is mainly accessed through a web portal and service oriented architectures based on web service technologies. The advantages of SaaS solutions are simplicity of integration, cost and scalability. The disadvantages of SaaS solutions is the perception of security issues. Ex Facebook, YouTube, Zomato etc.

B. PaaS (platform as a service)

PaaS comprises the environment for developing and provisioning cloud applications. The main users of this layer are developers wanting to develop and run a cloud application for a particular platform. The hardware and software within a PaaS solution is managed by the platform provider. Well known PaaS solutions providers include Windows Azure and Google app engine.

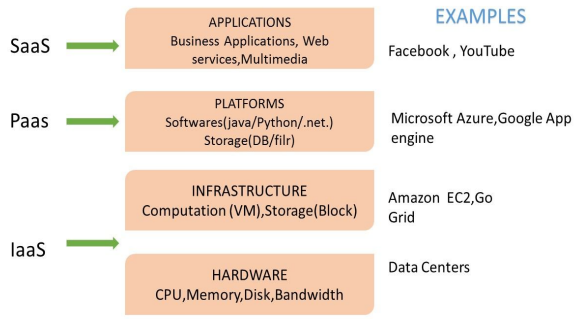

Fig 1-examples cloud service models

C. IaaS(Infrastructure as a Service)

The infrastructure as a service (IaaS) model provides a virtual data centre within the cloud. These resources are usually delivered as a virtualized platform and is not responsible for managing the underlying infrastructure. These important IT resources include services related to computing resources, the communications channel and data storage resources,. They assist existing applications to be provisioned on cloud resources and new services implemented on the higher layers. Within an IaaS solutions the developers must install their own operating system, support software and database management. The developers must manage both the hardware and software. Ex. amazon elastic compute cloud (ec [2]).

II. CLOUD DEPLOYMENT MODEL

a. Public Cloud

It is available for use by the general public. It may be owned by a large organization or company offering cloud services. As the cloud here is open it is less secure. A public cloud is generally the least expensive. Some of the best-known examples of public cloud systems are Amazon Web Services (AWS) containing the Elastic Compute Cloud (EC2) and the Simple Storage Service (S3) which form an IaaS cloud offering and the Google App Engine with provides a PaaS to its customers.

b. Private cloud.

It is owned by a specific entity and normally used only by that entity or one of its customers . The underlying technology may reside on or off site. A private cloud offers increase security at a greater cost. The chief advantage of these systems is that the enterprise retains full control over corporate data, system performance and security guidelines.

c. Community cloud

In a community cloud, organizations with similar requirements share a cloud infrastructure. It may be meant as a generalization of a private cloud and a private cloud. Community clouds are a form of clouds which are built and operated specifically for a particular group. These group have similar cloud requirements and their ultimate goal is to work together to achieve their business objectives. These are used by the organizations that work on joint projects, applications and ,building and executing..

d. Hybrid cloud

A cloud that consists of two or more private, public or community clouds. The hybrid model is also suitable for enterprises in which the transition to full outsourcing has already been completed, for instance, to combine community cloud services with public cloud services.

III. CLOUD COMPUTING: ITS ORIGIN AND EVOLUTION

The cloud computing started in the late 1980s with the concept of grid computing, for the first time, a large number of systems were applied to a single problem, that were scientific in nature and requiring exceptionally high levels of parallel computation.

In 2000,Amazon started a new product development strategy to provide cloud computing to external customers, and launched Amazon web service (AWS) on a utility computing basis in 2006.

On March 1, 2011, IBM announced the IBM framework to provide support for cloud computing. Among the various components of the framework cloud computing is a critical piece.

Oracle announced the Oracle Cloud on June 2012 and aspects of the Oracle Cloud are still being developed, this cloud offering is posed to be the first to provide users with access to an integrated set of IT solutions, including the Applications (SaaS), Platform (PaaS), and Infrastructure (IaaS) layers.

IV. VIRTUALIZATION

Virtualization is the use of software and hardware to create the perception that one or more entities exist although the the entities in actually, are not physically present. Using virtualization we can take one server appear to be many , desktop computer appear to be running multiple operating system simultaneously or a vast amount of disk space or drives to be available. The most common forms of virtualization include server virtualization, desktop virtualization , virtual networks, virtual storage. Figure 2 represents the schematic of virtualization processes used in Cloud computing.

A hypervisor or virtual machine monitor (VMM) is computer software, that creates and runs virtual machines. A hypervisor runs one or more virtual machines on a machine which is called as host machine. This machine can be a computer as well as a server. Each of the virtual machine is called a guest machine. The guest operating systems are represented by the hypervisor with a virtual operating platform. It manages the execution of the guest operating systems. There are various types of virtualizations are discussed below.

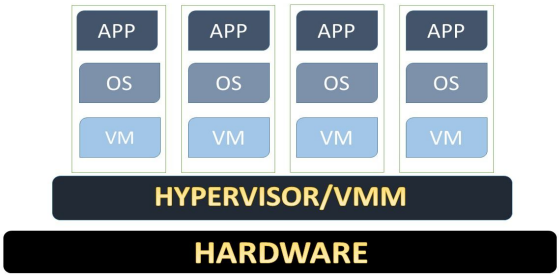

Fig 2- Virtualization process in cloud computing

a. Server virtualization

Making one server appear as many .Each virtual server may run the same or different operating systems. In order to decrease the server idle time a single physical server is virtualized to form multiple virtual servers. It can be concluded that the CPU utilization is the main reason for server virtualization.

b. Desktop virtualization

This allows to switch between various operating on the same operating systems, which makes the task easier for software developers and other tester staffs. This reduces the need for duplicate hardware and has other economical aspects.

c. Virtual networks

These create a illusion that a user is connected directly to a company network and resources, although no physical connection may exist. Virtual networks are sometimes called VPN (virtual private network),using a virtual private network the users can connect to a network and access the resources from any internet connected network.

IV. VIRTUALIZATION : ITS NEED

We have been hearing a lot about Virtualization as we talk about cloud computing. Most of the businesses often use a combination of a number of application servers, catalog server, web servers, image servers, file servers, video and audio servers, and the most important the database servers. Although modern web usage developments may suggest that all of the above mentioned hardware infrastructure is being used well almost all the time, this is falsehood and more accurately, an illogical belief. If 75% of the hardware appears to be used at any time it is considered to be underutilized. The servers typically take only about (1-10) milliseconds to service each request. Generally ,the amount of time the server machine is kept up and running relative to the actual time spent by it servicing the requests, is much higher .This clearly shows that a significant amount of energy is wasted per server in the process of keeping the servers up and ever-ready to service requests upon their arrival.

So the efforts to maximize the server utilization is limited by the number of incoming server requests. For ensuring that a good fraction of time is spent by the server in servicing requests, virtualization must be ensured

Virtualization technique ensures the availability of hardware and gives every application running on top of it. The details of the virtual, simulated environment are kept transparent from the application. The advantage here is the reduced cost of maintenance and reduced energy wastage which is not very surprising. So virtualization reduces the number of physical servers as a result of which one needs to maintain few servers ,this becomes much cheaper and easier.

The amount of energy wasted is a function of the number of physical servers that is reduced in a virtualized environment.

In case of desktop virtualization updates may now be made available much sooner as a single firmware update does not update one client machine, but several instances of the same.

V. VIRTUALIZATION AND MULTIPLE PROGRAMMING

In case of multiprogramming CPU is shared among processes. The Memory in multiple programming is shared using Page Tables. Process knows it is being managed- uses system calls. In Virtualization CPU is shared among OSs. Memory here is shared using more level of indirections. It includes multiple Page tables .OS may or may not know that it is being managed.

VI. DATA CENTRES

Data centres form the basis of a wide variety of services offered through the Internet including Web-hosting, e-commerce and general services such as software as a service (SAAS), platform as a service (PAAS), and grid/cloud computing . They consist of concentrated equipment to perform different functions like Store, manage, process, and exchange digital data and information. Data centres are mostly run by large companies or government agencies. This is a simple cage or rack of equipment for some and a room consisting of few cabinets for others. There is continuous increase in the demand for both the physical infrastructure and IT equipment due to rapid growth of data centres. This results in continuous increase in energy consumption. Computers require electricity, as well as protection from theft ie the accidental or intentional manipulation of hardware. So, one has to safeguard data centres against external influences and provide them with sufficient cooling. This increased consumption of energy causes an increase in the production of greenhouse gases which are hazardous for environmental health. Figure 3 indicates the city wise distribution of Data centres which shows Bangalore city is leading with 16% of data centres in India followed by New Delhi (12%) and Hyderabad (9%). It is also seen that almost all the direction in India there are many data centres.

Virtualization technology provides the solution but it has many overheads, like total cost of ownership, energy and efficiency calculations and return of investment done. The problem faced by IT managers is proper implementation of virtualization technology in data centres This paper comes up with a model to be followed by IT managers to properly implement virtualization in their data centres to achieve efficiency and reduce carbon footprints.

We have to consider a numerous things before virtualization, one is the architecture itself. It starts with the application. It has to be checked if the application can be deployed in a manner so that it can be virtualized and if it supports clustering or are there tools that help it support clustering so that each application instance recognizes state.In this case, that application is a great candidate for virtualization within the broader context of the application delivery network framework. If you can't replicate the data in real time, there might still be an opportunity to virtualize redundant sites if the data being served doesn't require up-to-the-minute freshness. There are a lot of scenarios where that does makes sense. Ultimately, you have to look at the underlying application infrastructure to determine what you can virtualize. The same is true for virtualizing connectivity and links. You also have to consider the amount of data and performance during the duplication process. The challenge is how much of that data can be concurrently transferred or put into the pipe while eliminating protocol communication overhead., which includes available bandwidth

.

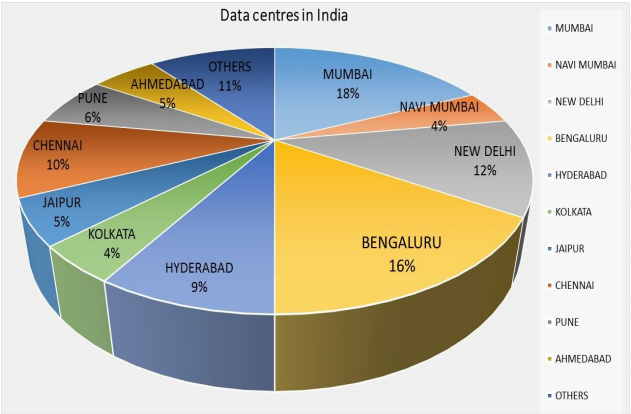

Fig 3 – Data centres in India. (Source-: datacentermap.com)

VII. VIRTUALIZATION PROCESS FOR X86 ARCHITECTURE

In x86 machine virtualization, a virtualization layer is added between the hardware and operating system. This virtualization layer helps in running multiple operating system simultaneously within the virtual machines in a single computer. This layer dynamically divides and shares the available resources basically physical such as storage, CPU, memory etc. .Servers can now run in tremendously fault tolerant configurations on virtual infrastructure 24x7x365 with no downtime needed for backups or hardware maintenance. The virtualization layer is mounted by the hypervisor layer on a clean x86-based system. Hypervisor has direct access to the hardware resources rather than going through an operating system. It is more efficient and delivers greater strength, scalability and performance. The functionality of the hypervisor varies greatly grounded on architecture and implementation. Each VMM that runs on the hypervisor implements the virtual machine hardware abstraction and is responsible for running a guest OS. Each VMM requires to be partition and share the CPU, memory and I/O devices to successfully and efficiently virtualize the system.

a. CPU Virtualization

The x86 operating systems are designed in such a way that they run on the bare-hardware and they assume that they possess the computer hardware. The x86 architecture provides four levels of privilege known as layer 0, 1, 2 and 3 to operating systems and applications to work across computer hardware. The user level applications and processes generally run in the last layer and the operating system has direct access to hardware and memory and must execute its privileged instructions in Layer 0. For virtualizing it is required to place a virtualization layer just below the operating system. This layer also has to manage the virtual machines. The situation gets complicated when some sensitive instructions cannot be virtualized properly. The main cause of this is that they have different semantics when they are not executed in layer 0.This problem was solved when binary translation was developed. This allowed the VMM to run in layer 0 and moving the operating system to a user level layer with less privilege than the VMM in Layer 0.

The three alternative techniques now exist for handling sensitive and privileged
instructions to virtualize the CPU on the x86 architecture:

- Full virtualization by binary translation
- OS supported virtualization or paravirtualization
- Hardware supported virtualization

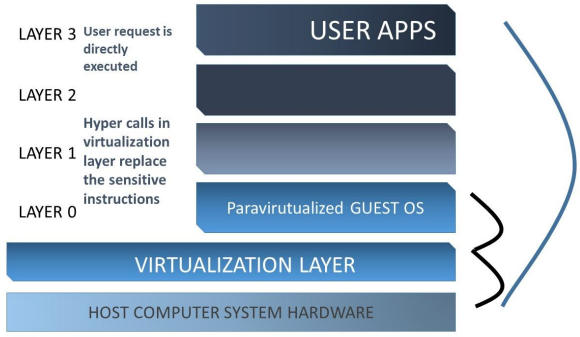

Fig 4 – Work flow of full virtualization

1) Full virtualization by Binary Translation: The kernel codes are translated using binary translation and direct execution techniques in order to replace the non-virtualized instructions with a sequence of instructions that can affect the hardware. During this time, the processor directly executes the user level code for high performance virtualization. Each of the virtual machine gets the services of the physical system which includes a virtual BIOS, virtual devices and virtualized memory management from the virtual machine monitor. The virtualization layer completely decouples the guest OS because of Full Virtualization offered by the combination of binary translation and direct execution. The guest OS is not aware it is being virtualized and needs no modification. Full virtualization is the option that needs no hardware assist or operating system to virtualize sensitive and privileged instructions. The hypervisor translates all operating system instructions and caches the results for future use. Full virtualization provides the best security and isolation for virtual machines, and provides easy migration as the same guest OS instance can run virtualized or on native hardware. Figure 4 presents the full virtualization work flow at different layers.

2) OS assisted virtualization or para virtualization: Para virtualization modifies the OS kernel in order to replace the non-virtualizable instructions with hyper calls which can communicate directly with the virtualization layer i.e.

hypervisor. It improves performance and efficiency, by communication within the guest OS and the hypervisor. The hypervisor also provides hyper call interfaces for other critical kernel operations such as memory management, interrupt handling etc. Para virtualization is dissimilar from full virtualization, as the unchanged OS does not know it is virtualized and sensitive OS calls are trapped using binary translation. As para virtualization requires deep kernel modifications, it can introduce support and maintainability issues in production environment. This virtualizes the processor and memory using a modified kernel and virtualizes the I/O using custom guest OS device drivers. It is easier than full virtualization which requires binary translation ie more difficult. The schematic of paravirtualization presented in figure 5.

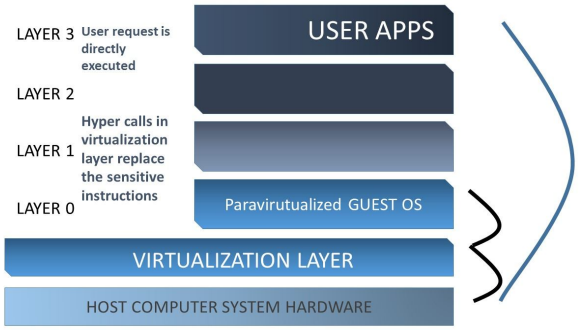

Fig-5 schematic of paravirtualization process

3) Hardware Assisted Virtualization: Hardware vendors have started embracing virtualization and developing new features to simplify virtualization techniques. The First generation developments include Intel Virtualization Technology (VT-x) and AMD's AMD-V that both target privileged instructions with a new CPU execution mode feature that allows the VMM to run in a new root mode below layer 0. The privileged and sensitive calls are set to automatically trap to the hypervisor and removes the requirement for either binary translation or paravirtualization. Virtual Machine Control Structures stores the guest state. (VT-x) or Virtual Machine Control Blocks (AMD-V).The processors with above configuration were made available recently and only very few and new system contain such features that assist the hardware.

b. Memory Virtualization

The next critical component is memory virtualization. This includes the sharing the physical system memory and dynamically allocating it to virtual machines. A contiguous address space is seen by applications the that is not essentially tied to the underlying physical memory in the system. The operating system maps the virtual page numbers to physical page numbers that are stored in page tables. All modern x86 CPUs contain a memory management unit (MMU) and a translation look aside buffer (TLB) for improvement virtual memory . In order to run multiple virtual machines on a system, another level of memory virtualization is essential. Thus, one has to virtualize the MMU to support the guest OS. The guest OS controls the mapping of virtual addresses to the guest memory physical addresses, but the guest OS cannot have access to the actual machine memory. The TLB hardware is used by the VMM to map virtual memory directly to the machine memory in order to to avoid the two levels of translation on every access. The guest OS changes the virtual memory to physical memory mapping, during this time the VMM updates the shadow page tables to enable a direct lookup. Some overhead for all virtualization approaches is created by the MMU virtualization . in this area second generation hardware assisted virtualization will offer efficiency gains

VII. ECONOMIC ASPECTS OF VIRTUALIZATION

In the above, we saw the process of virtualization process reduces the requirement of physical machines like number of computers, servers storage amount etc. The companies which provide data centres use this process for storing data, softwares etc. which in turn increases the profit of the companies. This concept reduces the requirements of large numbers of servers and associated parameters like electricity, cooling costs etc. which results benefit for the clients also as there will be huge cost reduction in the cloud usage.

VIII. RELATED WORK

A number of work have been carried out in the basic concept of cloud computing including the various models of cloud. [1-4]. The need of virtualization and the basic concepts of it are well explained in the work],[5-7]. The work explains the proper importance of data centres [8].,but it doesnot contain the importance of virtualization in data centres . In another work [9] the authors explains the model of implementing the virtualization .

IX. CONCLUSIONS

Using Cloud Computing the complexity and cost of owning and operating computers and networks can be significantly reduced. Customization of Cloud services can be done and it is also flexible to use, advanced services can be offered by the providers so that an individual company might not have to spend money or expertise to develop. In this work cloud computing and virtualization has been briefly introduced. This contains the evolution of cloud computing. The cloud models are explained with proper examples. The need of virtualization and the process involved in it is depicted. The advantages of virtualization are given with brief explanation. The requirement and importance of virtualization in data centres have also illustrated. How virtualization reduced the problems of machines has also been depicted. The virtualization techniques adopted in x86 architecture is shown

here. The economic aspects of virtualization are also been discussed.

ACKNOWLEDGEMENT

The two authors Anurag and Dines acknowledge to SPARK program of CSIR C-MMACS for providing the internship, HOD, Dept. of Computer Science and Engineering and Director, National Institute of Technology, Rourkela for providing the necessary infrastructures to carry out the work.

REFERENCES

[1] Gouda K C, Radhika T V, Akshatha M, Priority based resource allocation model for cloud computing, International Journal of Science, Engineering and Technology Research (IJSETR), ISSN: 2278 – 7798, pp 215-219.,2013.

[2] Gouda, K. C., and Dwaipayan Acharjee. "An Outlook of cloud computing." International Journal of Science, Engineering and Technology Research 2.6 (2013): pp-1403-1406.,2013.

[3] Jamunarani M R., Gouda, K. C., & Nirmala, H. Load Balancing Technique for Climate Data Analysis in Cloud Computing Environment, International Journal of Computer Science, Engineering and Technology, 3(5) (2013): pp-183-185.,2013..

[4] K C Gouda , Sambeet Kar , Rohan Patra."Cloud Computing: What, How and Why". International Journal of Engineering Trends and Technology (IJETT). V4(6):2453-2457 Jun 2013. ISSN:2231-5381,2013 .

[5] Norman Wilde and Thomas Huber ,Virtualization and Cloud Computing' , University Of West Florida.

[6] Intel Virtualization Technology (VT) Explained', available at http://www.hardwaresecrets.com/printpage/263

[7] "Virtualization and Cloud Computing Steps in the Evolution from Virtualization to Private Cloud" by intel IT centre

[8] Mauricio Arregoces, Maurizio Portolani , Data centre fundamentals"

[9] Introduction to virtualization: Get started with ESXi', Vmware.