A Survey on Virtualization and Hypervisor-based Technology in Cloud Computing Environment

Sonam Srivastava
M.tech Computer Science and Engineering
Madan Mohan Malaviya University of Technology
Gorakhpur, Uttar Pradesh, India

S.P Singh
Associate Professor (CSE dept.)
Madan Mohan Malaviya University of Technology
Gorakhpur, Uttar Pradesh, India

***Abstract*— Cloud computing is one of the today's largest hearing fields and exciting technologies, because it is flexible and scalable, also it reduces the cost and complexity of applications. The core technology that adds to the features of cloud computing is virtualization. Virtualization is a component of cloud computing. On the basis of virtualization cloud computing allows easy deployment of workloads and their quick scaling through the rapid provisioning of virtual or physical machines. In this paper we have discussed basic virtualization technology and also several reasons as to why it has gained immense impetus in the recent years. We have also compared the scenario before and after the advent of virtualization. Virtualization works at several levels thus, in this paper we present a view on each of these levels. We also discuss hardware virtualization techniques at the hardware level which distinguish from each other on the basis of the kind of support they expect from the hardware underneath and whether the guest should be modified or not. The entire internal working of virtualization layer is dealt by the hypervisor. Hypervisor is the fundamental element of Hardware level virtualization. Furthermore, in this paper we indicate that there exist two types of hypervisors and also discuss their security concerns along with the internal organization of the hypervisors. We also present a view on certain open issues and research directions that the researchers can consider and exploit further.**

Keywords—cloud computing; virtualization; operating system; hypervisor

I. INTRODUCTION

Cloud computing is the most popular and the budding concept among all the other computing paradigms such as peer to peer computing, grid computing, traditional computing, distributed computing and network computing. It is a novel paradigm that offers computing as a service on demand. Cloud computing paradigm allows the users to access these services irrespective of how they are hosted and delivered. Data centres considered as pool of computing services which deliver the services based on pay-per-use model where the user enjoys the benefits of cost saving, scalability, flexibility and elasticity to name a few. Cloud computing is economical as the users access the service as per their requirements and the amount is incurred only when they are in use. Cloud computing is an Internet-based emerging technology which aims at storing and processing deluge of data. On the basis of virtualization cloud computing allows easy deployment of workloads and their quick scaling through the rapid provisioning of virtual or physical machines. Various authorized users can have access to the data and records hosted on the clouds from anywhere across the world with the help of any device. Several distributed computers efficiently work together to process huge amounts of data, while assuring the brevity of processing time of query results to users. All the entities in the cloud are linked together. If one of them is succumbed to failure or crashes the data is still available for others to use.

The four deployment models of cloud computing include private, public, community and hybrid cloud deployment models. The most commonly used model is the public model. These are owned and operated by a third party service provider. However there are certain security concerns associated with the model, but it is rendered best model for start-ups, testing and when heavy loads are to be handled. Service of a private cloud is managed by a particular organisation and there are no security concerns. Hybrid cloud unites the advantages and features of both private and public clouds.

Clouds are considered as a pool of easily usable and accessible resources virtualized resources. The three cloud service models are Software as a service model, Platform as a service model and Infrastructure as a service model. Software as a service is popular service development model that ensures that users pay for using software not for owning it. Platform as a service model provides the users with a platform to create their own software and configure it according to them. The self service model that manages and monitors the remote data is Infrastructure as a service model.

Cloud provider manages the storage, computing and networking services. Cloud computing is basically accessing services that are needed to performing various activities with dynamically changing users demands. The idea behind cloud computing is to enable businesses by increasing efficiency, resource utilization and flexibility of their computer hardware. The key technology that makes it possible is virtualization. Virtualization plays an important role in cloud computing, since it allows for appropriate degree of customization, security, isolation and manageability that are fundamental for delivering services on demand.

ISSN: 2278 – 1323

II. Virtualization in detail

A. *Basic virtualized technology and architecture*

Virtualization is the backbone of cloud computing. It is one of those fundamental components of cloud computing that allows creation of isolated, customizable and secure execution environment for running different applications without affecting other users' application. It is a vast umbrella of technologies that are meant to provide an abstract environment whether virtual hardware or an operating system to run applications. Virtualization basically is creating an emulated version of something. It also allows multiple operating systems to be able to run on the same machine. Virtualization architecture comprises of four layers, the topmost layer is the application layer, and then comes the operating system layer on top of virtualization layer (Hypervisor) and the hardware layer being the bottommost layer. Figure 1 shows the four layer virtualization architecture. The two intermediate layers offer easier interface to the applications above. Hypervisor is a software program that manages the multiple operating systems to be able to run on the same machine. It also monitors system processors memory and other resources in order to be able to able to allocate it according to each processor requirements. Hypervisor is much more efficient than other hosted architectures enabling greater performance, robustness and scalability.

Virtualization assures complete separation of applications from the operating system from the underlying hardware. The most commonly available virtualization software is VMware. This imitates hardware resources of one physical machine to create fully functional Virtual machine. An operating system and associated applications can be installed on this machine similar to as done for the physical hardware. An operating system on a virtual machine is called a guest operating system. Now, various instances of operating system can be created on a same physical hardware as separate entities. Hypervisors will assign their own resources to each guest operating system. Therefore, there is no interaction between processes running on different guest operating system and also there is no communication between them using regular operating system primitives of semaphores, shared memory, pipes or signals.

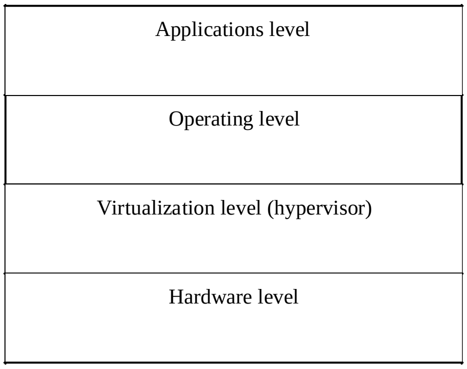

Figure 1: Virtualization architecture

B. *Reason for the immense importance of virtualization technology*

In the recent years virtualization technologies have gained renewed interests due to its various effects and benefits.

- Isolates different operating systems along with their applications from others. Their separation results in the elimination of any chance of interference between the activities of operating systems.
- It fosters better resource utilization and causes increase in performance. Major portion of the capacity of desktop computers is rarely utilized, which otherwise fulfils all the needs of everyday computing. So, this unutilized space and resources can be used to host virtual machine manager and execute distinct virtual machines with by far acceptable performance.
- Virtualization renders fault tolerant environment. That is whenever a particular server suffers from power failure then that physical hardware dies. However, all the email services are still available for client computers.
- Running processes can be migrated from one physical machine to another physical machine. This process migration brings about transparent movement to processes from highly loaded machines to less loaded ones. Thus, it is helpful in load balancing and saving energy. Process migration does not cause any disruption of activity during migration.
- Data centres are considered as a pool of computing resources. The increasing demand for storage and additional capacity has led to increase in number of data centres. A lot of energy is required in keeping them cool as they intake a huge amount of power. Various IT enterprises have come up with greening initiatives of adopting virtualization so that through server consolidating it can provide them with the solution to limit the impact that the company has on the environment.
- Virtualization decreases the number of needed servers for a given workload and thus reducing the cost of administrative personnel through server consolidation. It also resolves the resource utilization and reduces power consumption of data centres by aggregating multiple servers and applications originally deployed on different servers on one physical server.
- Virtualization offers a stable and convenient way to create a reproducible environment for software testing. This provides an easy way to test and debug new software before and during deployment into production environments.

C. *Traditional server vs. virtual server*

Servers are one of the very needful hardware components in cloud computing paradigm. Responding to several request propagated by clients and balancing load across different components of network are the purposes served by servers.

Traditional versus virtual servers depict the scenario before and after the advent of virtualization.

- Traditional server: Traditional server is basically a three layered architecture. It consists of hardware layer as the bottommost layer followed by operating system layer on top of it and application layer being the topmost layer. Traditional servers are normally maintained by the system administrators. We can easily deploy things with the help of traditional servers. Here, only single operating system could be present on the physical hardware. With traditional servers the migration of operating system from original hardware to other was time consuming and was a tedious job. This also incurred heavy cost. If multiple operating systems are run on the same machine it would create conflicts.

- Virtual server: Virtual server model came into existence as a solution to the problems of traditional server. Virtual server is a four layered architecture. The four layers include hardware layer, virtualization layer, operating system layer and application layer starting from bottom level to topmost level. Virtual server seeks to encapsulate the server software from the server hardware. The server configurations can be changed while the services are running. Business endeavours are rendered more flexible with the emergence of virtual server model. Virtualization layer stands above the hardware layer, users can easily transfer the operating system by just performing copy and paste operation to the new physical hardware. The application associated with operating system stays intact with it.

Table 1: Comparison

Traditional server	**Virtual server**
Scenario before the advent of virtualization.	Scenario after the advent of virtualization.
Three layered architecture.	Four layered architecture.
Only single operating system can be installed on physical machine.	Virtualization allows multiple operating systems to be installed on the same physical machine.
No separation between the hardware and the software.	Complete separation between the hardware and software.
Migrating the only operating system from the original physical hardware to another is a tedious job.	This model allows software portability from one physical machine to another.
Infrastructure setup incurs huge amount of cost and is inflexible too.	Infrastructure setup is cost effective.
Operating system and its applications are hardware dependant.	Operating system and its applications are independent of its hardware.
There is no chance of physical infrastructure up gradation possible.	Infrastructure up gradation can be done on timely basis.
Virtualization layer does not exist in this model.	All the working of the virtual server model is reliant on the virtualization layer internally handled by Hypervisors.
Existence of resource underutilization.	Resources are rendered properly utilized.

III. Levels of virtualization

Virtualization is mainly used to imitate execution environment, networks and storage. The oldest, most popular and with the most developed area among these categories is the execution virtualization. Execution virtualization can be further dissected into two major categories by considering the host type they require. The two categories include process level and system level techniques. Process level techniques are implemented on top of existing operating system which has full control of the hardware. The system level techniques are implemented directly on top of the hardware and they do not require any support from the existing operating system. Operating system level virtualization, programming language level virtualization and application level virtualization are the three categorizations of process level techniques. Hardware assisted virtualization, full virtualization, paravirtualization and partial virtualization are the four categorizations of hardware virtualization techniques at the hardware level which comes under system level techniques.

A. Operating system level virtualization

The opportunity to create a separated execution environment for certain applications that are managed concurrently is provided by operating system level virtualization. The fact that there is no hypervisor or virtual machine manager and also virtualization is done within a single operating system differentiates it from hardware level virtualization. This strategy is an efficient solution for server consolidation and it also imposes little or no overhead as the applications can directly use operating system calls. Examples of OS level virtualization are: openVZ, iCore Virtual accounts, free virtual private servers (free VPS), IBM Logical Partition (LPAR).

B. Programming language level virtualization

Ease of deployment of applications, portability across different operating system and platforms and managed execution are certain features rendered by Programming language level virtualization. Nowadays, the most popular

ISSN: 2278 – 1323

technologies for application development are represented by java platform and the .NET framework. The major goals of this level of virtualization include providing uniform execution environment across different platforms and it also allows for more control over the execution of various programs since direct access to the memory is restricted.

C. Application level virtualization

Application level virtualization works only for a specific environment as it provides support for all applications running on top of specific environment. In this scenario applications run as if they are installed on a specific runtime environment, although they are not. Application level virtualization uses either of the two strategies, interpretation or binary translation to execute program binaries compiled for different hardware architectures. Wine is the most popular solution implementing application virtualization.

D. Hardware virtualization techniques at hardware level

- Hardware assisted virtualization depicts the scenario where the virtual machine manager able to run a guest in complete isolation, is build by architectural support provided by the hardware.
- Full virtualization is the ability to run a program, most likely an operating system on top of a virtual machine directly and without any modification. Complete isolation is the key advantage of full virtualization. It leads to coexistence of distinctive system on same platform, enhanced security and ease of emulation of architecture. It combines the hardware and software when together to prevent the harmful instructions to be executed directly on the host, leading to an efficient and sufficient implementation of full virtualization.
- Paravirtualization allows simpler and easy implementation of virtual machine managers. It exposes a software interface to virtual machine that is to a very little extent modified by the host, and as a result of which guest also needs to be modified. The main goal is to prevent performance losses experienced during managed execution, by demanding performance critical operation to be executed directly on the host.
- Partial virtualization restricts the complete execution of the guest operating system in complete isolation, by providing partial emulation of underlying hardware. It is an essential tool for achieving full virtualization.

E. Storage virtualization

Storage virtualization if adopted by the users they do not have to be concerned about the exact location of their data or records and logical path can be used to identify them. Wide range storage facilities can be harnessed and represented under a single logical file system. One of the most popular techniques for storage virtualization is storage area networks (SANs).

F. Network virtualization

Network virtualization unifies hardware appliances and specific software for the management and creation of a virtual network. It can also aggregate several physical networks into a single logical network (external network virtualization), or provide a network like functionality to an operating system partition (internal network virtualization). The result of an external network virtualization is a virtual LAN (VLAN).

IV. Virtual machine manager or hypervisor

A. Basic idea and internal organisation of hypervisor

There is basically two type virtualization software available namely the client virtualization software and the hypervisor. Client virtualization software is installed on the normal operating systems such as Windows or Linux that are installed on the physical hardware. This enables the creation of as many instances of the already existing operating system. New operating system that is created is called an instance. All these newly created instances get installed on the existing operating system. They can also be migrated by just copying and pasting them to whichever server they need to be transferred.

Fundamental element of hardware level virtualization is virtual machine manager (VMM) or the hypervisor. In hardware level virtualization model the host is represented by physical computer hardware, guest by the operating system virtual machine by its emulation and virtual machine manager by the Hypervisor. The entire working of virtualization layer is internally handled by the hypervisor. It is a software program that allows the abstraction of underlying hardware. It brings about the recreation of the environment where the guest operating systems are installed. ESXi is the most popular hypervisor provided by the VMware. It is installed directly on the top the physical hardware on which the virtual environment is to be created. It follows the normal installation process. Hypervisor is managed via management software called Vsphere. This is installed on the computer used to administer the hypervisor. This Vsphere connects with hypervisor to create as many virtual computers in the virtual environment. Above this are installed any normal operating systems namely Linux or Windows. Hypervisor is the more powerful virtualization software as compared client virtualization software. It is not too complicated to use. Also it provides much more fault tolerance than the client virtualization software. ESXi hypervisor is available free of cost whereas users have pay for the management software. Hypervisor controls the hardware and this capability of it allows hypervisor based systems to have secure infrastructure. Type 1 and Type 2 are the two major categories of hypervisors.

- Type 1 hypervisors run immediately on top of the hardware hence they are called native virtual machine. That is they actually occupy the place of the operating systems and thus they can easily interact with the instruction set architecture interface exposed by hardware below. This type of hypervisors emulates the ISA interface in order to manage the guest operating systems.
- Type 2 hypervisors are hosted within an operating system hence they are called hosted virtual machine. They provide the virtualization services under the support of this operating system. This type of

hypervisors interacts with the operating system with the help of application binary interface and imitates the ISA of the underlying hardware for guest operating systems.

Fig. 1 shows the Hypervisor reference architecture. Internal organization of hypervisors comprises of three basic modules namely dispatcher, allocator and the interpreter. The entry point of the monitor is the dispatcher and it basically reroutes the instructions that are issued by virtual machine instances to the other modules. The responsibility of the allocator is to decide which system resources should be allocated to the virtual machines. The interpreter consists of interpreter routines which are executed whenever virtual machine executes a privileged instruction. Goldberg and Popek established the established criteria that need to be met by VMM to efficiently support virtualization are: guest running on the VMM should exhibit the same behaviour as when run on the physical host, virtualization resources should have complete control over the VMM and VMM should not interfere with the execution of dominant fraction of machine instructions.

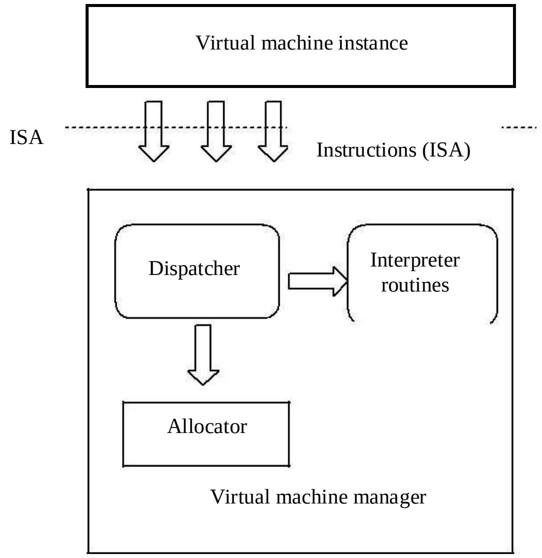

Figure 2. Hypervisor reference architecture

B. *Hypervisor security (benefits and downsides)*

- Authentication, authorization and networking are the three major levels of security management of hypervisors.
- Hypervisor is used as an abstraction to separate the virtual environment from the hardware below. This isolation prevents interference of any activities between them.
- Hypervisor simplifies the transaction management process in the cloud environment thus controlling all the access between the guests' operating system and shared hardware underneath.
- Since, the hypervisor lies below the guest operating system in virtualization hierarchy it easily detects any attack if it passes through the security systems in guest OS.
- In a hardware level virtualization there is only a single hypervisor, that is in a system exists a single point of failure. If hypervisor crashes or goes down due reasons of overload or successful attacks, entire system and virtual machines would be affected.
- As compared to other technologies the hypervisor also has vulnerabilities to certain attacks such as buffer overflow.

V. Virtualization downsides: limitations

There are certain limitations exhibited by virtualization technology. The most evident is represented as decrease in performance of guest operating systems, as a result of intermediation performed by virtualization layer.

- Due to interposition of abstraction layer in the midst of guest and the host, guest can experience increased latencies and delays. In case of Hardware virtualization when the bare metal is emulated on top of which entire system can be installed, performance degradation is caused by the overhead introduced by the following activities: support of paging with virtual machine, console functions, maintain the status of virtual machines, support privileged instructions.
- Sometimes due to virtualization host is rendered inefficient. There are certain features of host that are not exposed by the abstraction layer and hence they become inaccessible. In case of Hardware virtualization this could happen when for device drivers where the default graphic card provided by virtual machines maps only a limited subset of features available in the host. In case of programming level certain features of the underlying operating system are not accessible unless and until certain specific libraries are used.
- There are certain threats security holes associated with the virtualization. It opens the door to an unexpected and a novice form of phishing. The ability of complete and transparent emulation of host has led the way to certain programs designed specially to extract useful and sensitive information from the guest.
- Since, virtual machines have to share their data and interact and communicate with each other. So, if these communications do not meet considerable security parameters then they have the potential of becoming attacks target.

VI. Open issues and research directions

Certain downsides of virtualization and security concerns of the hypervisor can be exploited further to define new research directions. Generally, the Infrastructure as a service cloud providers give the illusion of frictionless registration process to their customers and also of unlimited storage and network capacity. This allows anyone to get started using the service. The solution to ensure security of data stored in the cloud is encrypting it against malicious attacks. However, there are certain other issues in this context such as ensuring integrity of data. When malicious and non-malicious users compromise the integrity of users' data there is no such mechanism available with the client to analyse its integrity of original data. Hence, new technologies must be adopted in order to preserve the integrity of users' data.

Another issue relates to benefits of storage virtualization in future network architectures. It is the fundamental building block of network paradigms and algorithms even then concept of storage is not fully defined in the networks. Immediately as the data reaches the network, backup should be prepared for it. Simply bootstrapping individual elements of the network from configuration files may not bring back the entire network functionality. Network virtualization can be exploited more in this case. Operating system virtualization is used to handle numerous administrative domains above the virtualization layer and hardware virtualization is to handle the same beneath the virtualization layer. The coexistence of multiple virtualization techniques in different administrative domains can be explored further in order to realize the benefits of end to end virtualization. Since, virtualization optimizes the cloud computing idea and enables creation and debugging of a new instance of an operating system on the same physical hardware. This allows for new research ideas to be implemented.

VII. Conclusion

In the above sections we describe the basics of cloud computing along with its relationship with virtualization. In the next section virtualization in detail and also its basic technology, reason for the immense importance of virtualization in the recent years has been highlighted. We have also compared the virtual server model with the traditional server which depicts the scenario before and after the advent of virtualization. It is a technology that that powers cloud computing. The various levels at which virtualization works have been discussed in the next section, with each level discussing its goals and how it serves as an essential part of virtualization as a whole. Moving on to the next section we deal with the fundamental element of virtualization technology the VMM popularly known as the hypervisor. The internal working of virtualization layer is entirely managed by the hypervisor. Virtualization is the backbone of cloud computing, without which it would be rendered an inefficient technology. However, every coin has two sides so in case also we have also discussed the other side of it that is the certain downsides of virtualization. This can be exploited further and taken into consideration for future research directions. Since, virtualization optimizes the cloud computing idea it has attracted wide attention in the recent years.

Acknowledgment

The completion of this paper would be incomplete unless we express our deep gratitude to the authors whose work we have taken for reference. Their papers and books have provided us with constant guidance and encouragement without which the successful completion of this paper could not have been made possible.

References

[1] Shyam Patidar, Dheeraj Rane and Pritesh Jain "A survey paper on cloud computing" 2012 Second International Conference on Advanced Computing and Communication Technologies on 7-8 January 2012 pp. 394-398 ISBN: 978-1-4673-0471-9 DOI: 10.1109/ACCT.2012.15 2012 IEEE.

[2] Mutum Zico Meetei "Cloud computing and security measure" 2013 Sixth International Congress on Image and Signal Processing (CISP) on 16-18 December 2013 pp. 852-857 ISBN: 978-1-4799-2763-0 DOI: 10.1109/CISP.2013.6745284 2013 IEEE.

[3] Farzad Sabahi "Secure virtualization for cloud environment using hypervisor-based technology" International Journal of Machine Learning and Cloud Computing, Vol. 2, No. 1, February 2012.

[4] Zongzian He and Guangqing Liang "Research and evaluation of network virtualization in cloud computing environment" 2012 Third International Conference on Networking and Distributed Computing on 21-24 October 2012 pp. 40-44 ISBN: 978-1-4673-2858-6 DOI: 10.1109/ICNDC.2012.18 2012 IEEE.

[5] Jun Huang, Yanbing Liu and Qiang Duang "Service provisioning in hypervisor-based cloud computing: Modelling and Optimization " Global Communications Conference (GLOBECOM) on 3-7 December 2012 pp. 1710-1715 ISBN: 978 -1-4673-0920-2 DOI: 10.1109/GLOCOM.2012.6503361 2012 IEEE.

[6] Krishna Tej Koganti, Eshwar Patnala, Sai Sagar Narasingu and J.N Chaitanya "Virtualization technology in cloud computing environment" International journal of Emerging Technology and Advanced Engineering Vol. 3, Issue 3, March 2013

[7] Ashiq khan, Alf Zugenmaier, Dan Jurca and Wolfgang Kellerer "Network virtualization: A hypervisor for the Internet" Communications Magazine, Vol. 50, Issue 1, January 2012 pp. 136-143 DOI: 10.1109/MCOM.2012.6122544 2012 IEEE.

[8] Xiangyang Luo, Lin Yang, Linru Ma, Shanming Chu and Hao Dai "Virtualization security risks and solutions of cloud computing via divide and conquer strategy" 2011 Third International Conference on Multemedia Information Networking and Security(MINES) pp. 637-641 ISBN: 978-0-7695-4559-2 DOI: 10.1109/MINES.2011.54 2011 IEEE.

[9] Dr. Rajkumar Buyya, Dr. Christian Vecchiola and Dr. S Thamarai Selvi "Mastering cloud computing".

[10] Davide Mulfari, Antonio Celesti, Massimo Vellari and Antonio Puliafito "Using virtualization and noVNC to support assistive technology in cloud computing " 2014 Third Symposium on Network Cloud Computing and Applications on 5-7 February 2014 pp. 125-132 ISBN: 978-0-7695-5168-5 DOI: 10.1109/NCCA.2014.28 2014 IEEE.