Scientific Journal of Impact Factor (SJIF) : 3.134

ISSN (Print) : 2348-6406
ISSN (Online): 2348-4470

International Journal of Advance Engineering and Research Development

An Overlook of Coin Word Virtualization and Snapshot

Arvind Iyer[1], Prof Vatsal Shah [2]

[1]*Student ,Information Technology Department, Birla Vishvakarma Mahavidhalaya,ayalurarvind@gmail.com*
[2] *Assitant Professor,Information Technology Department, Birla Vishvakarma Mahavidhalaya, vatsal.shah@bvmengineering.ac.in*

Abstract : *Virtualization a fifteen letter Buzzword word that has amazed everyone which has a such a powerful application in computing world. The interest in virtualization has been growing steadily among the computing community in the last few years. Virtualization provides many benefits which most of scientist and research state as – greater efficiency in CPU utilization, greener IT with less power consumption, better management through central environment control, more availability, reduced project timelines by eliminating hardware procurement, more central control of the desktop, and improved outsourcing services etc and list goes on. With these benefits, it is no wonder that virtualization has quickly evolved from concept to reality. The technologies that are used for servers, clients, storage, and networks are being virtualized to create flexible and cost effective infrastructures. Snapshot a coin word root lies in photography which is related to photo click ,in the same way snapshot in computing world is a state of machine at a particular time. Virtualization is now a basic fundamental like programming language rather than concept so the focus of this review paper is to familiarize with the coin word Virtualization, it's types ,its application ,features and use of snapshot in virtualization.*

Keywords: *Virtualization,Snapshot,Memory,Hpervisor,Virtual machine,Hardware,Architecture*

I. INTRODUCTION

The aim of this article is to present the basic principles of virtualization and use of snapshot in virtualization .This article will have brief introduction to coin word virtualization ,its types,benefits and use of snapshot in virtualization.

Virtualization as a concept was developed for the first time by IBM in 1960s. During 1980s and 1990s, desktop computers and x86 servers became much more available and, therefore, virtualization had not undergone any significant progress at that period. The client-server applications and emergence of operating systems Windows and Linux made both the server technologies and home PCs markedly cheaper. However, new challenges have appeared, such as time-consuming maintenance of the existing computers, high operating cost or high failure rate, which led to the development of x86 virtualization. Virtualization has increased the effectiveness dramatically and reduced the overall IT expenses.[1].

In general virtualization is concept in which so called virtual objects are created. These objects are virtual which means not real .So virtualization is a non tangible concept .Virtualization can be performed at various levels of computer ranging from hardware to software .We can have virtual disk ,virtual memory ,virtual processor etc by use of virtualization tools. Virtualization allows more operating systems run in one physical machine, e.g. Windows, Linux etc. The number of virtual machines that can run simultaneously in one machine depends on hardware specification. At present, virtual machines are implemented in most data centers as normal servers but with much less maintenance and administration cost. This technology potential is huge and plays a great role in the computer science future [2]

In today's world online backup has become most important issue. Snapshot is a feasible method to ensure the proper persistence availability of data .In computer systems, a snapshot is the state of a system at a particular point in time. The term was coined as an analogy to that in photography. It can refer to an actual copy of the state of a system or to a capability provided by certain systems.[3] There are three basic type of snapshot,one is based on file system, another based on storage sub system and last one based on volume manager.

II. VIRTUALIZATION BENEFITS

Consolidation of servers definitely is the greatest benefit of virtualization. Virtualization has become a saving means of about 60% of companies. Their IT departments have been transforming the physical servers to the virtual ones. Such transformation has several advantages .

A). **Continuity of operation and high availability**

The virtual server is separated from a particular physical hardware and can easily be backed up and restored in another hardware or another locality. Advanced virtualization technologies thus allow the transfer of a virtual server to another physical server without disconnection of working users, even in applications not supporting the high availability solution at all.

B).**Scalability**
The separation of the physical hardware from the individual virtual servers and stations allows an easy increase of the applications performance by a simple addition of a new hardware in the existing infrastructure. If, for example, the number of demands made on a server is increased, the memory or CPU performance can simply be increased in the administration console. The same it is with the addition of any other hardware allowed by the given virtualization technology. Attendance simplification and cost of operation reduction Thanks to the reduction of the number of physical servers, automated installation of new servers from templates, their easy transferability to another hardware, and thanks to the centralized administration tools, the attendance activities during the servers and applications installation and operation have been simplified significantly.

C).**More effective administration**
The symbiosis of a company IT environment and its commercial needs should be the primary aim of the consolidation. The aim of the IT consolidation is mainly to shorten the period of the return on investment in the IT infrastructure and to reduce the expenses related to IT administration. The administration of an increasing number of servers, working stations and other devices of the company IT infrastructure can be expensive and demanding from the capacity point of view. Working stations consolidation eliminates the autonomous administration disadvantages to a great extent. IT departments can administer and maintain the software on such stations much more effectively, including the central application of safety patches and correction packets. A unique, consolidated and consistent environment in a company can increase the employees' labour productivity and reduce the expenses. A unique setting and compliance with the configuration policy guarantees much less vulnerability of the centrally administered stations towards the safety threats[4]

III. TYPES OF VIRTUALTIZATION

Generally speaking, virtualization falls into three categories: Operating System, Storage, and Applications.It depends on an person to person how he defines it and classifies it. Around the world wide web and different books one may find a huge content on the classification of virtualization. The best classification and introduction to each according to us of specific categories is as follows:

A).**Software (Operating System) Virtualization**
The most used out form of virtualization today, virtual operating systems (or virtual machines) are quickly becoming a core component of the IT fraternity. Generally, this is the form of virtualization end-users are most familiar now days by using software like vm worksation and virtual box etc. Virtual machines are typically full implementations of standard operating systems, such as Windows Vista or RedHat Enterprise Linux, running simultaneously on the same physical hardware. Virtual Machine Managers (VMMs) manage each virtual machine individually; each OS instance is unaware that 1) it's virtual and 2) that other virtual operating systems are (or may be) running at the same time. Companies like Microsoft, VMware, Intel, and AMD are leading the way in breaking the physical relationship between an operating system and its native hardware, extending this paradigm into the data center. The ultimate goal it is to save money of the enterprises on hardware, co-location fees, rack space, power, cable management, and more.

B).**Server Virtualization**
Application Server Virtualization has been around since the first load balancer, which explains why "application virtualization" is often used as a synonym for advanced load balancing. The core concept of application server virtualization is best seen with a reverse proxy load balancer: an appliance or service that provides access to many different application services transparently. In a typical deployment, a reverse proxy will host a virtual interface accessible to the end user on the "front end." On the "back end," the reverse proxy will load balance a number of different servers and applications such as a web server. The virtual interface—often referred to as a Virtual IP or VIP—is exposed to the outside world, represents itself as the actual web server, and manages the connections to and from the web server as needed. This enables the load balancer to manage multiple web servers or applications as a single instance, providing a more secure and robust topology than one allowing users direct access to individual web servers. Application Server Virtualization can be applied to many types of application deployments and architectures and it is gaining a huge popularity in research arena as it has many horizons to be covered as day to day limit of data and severs optimizing it are increasing.

C).**Network Virtualization**
Network virtualization is one of the most ambiguous, specific definition of virtualization is tough to describe as Network is wast domain and each has different saying on it. A simple example of network virtualization is a VLAN: a single Ethernet port may support multiple virtual connections from multiple IP addresses and networks, but they are virtually segmented using VLAN tags. Each virtual IP connection over this single physical port is independent and unaware of others' existence, but the switch is aware of each unique connection and manages each one independently. Another example is virtual routing tables. The single routing table will contain multiple routes for each virtual connection, but they are still managed in a single table. Virtual routing tables change that paradigm into a one:many relationship, where

any single physical interface can maintain multiple routing tables, each with multiple entries. This provides the interface with the ability to bring up (and tear down) routing services on the fly for one network without interrupting other services and routing tables on that same interface.

D)**.Hardware Virtualization**

Hardware virtualization is very similar in concept to OS/Platform virtualization, and to some extent or rathe saying dependent on OS virtualization will not be cruel. Hardware virtualization breaks up pieces and locations of physical hardware into independent segments and manages those segments as separate, individual components. Although they fall into different classifications, both symmetric and asymmetric multiprocessing are examples of hardware virtualization. In both instances, the process requesting CPU time isn't aware which processor it's going to run on; it just requests CPU time from the OS scheduler and the scheduler takes the responsibility of allocating processor time. As far as the process is concerned, it could be spread across any number of CPUs and any part of RAM, so long as it's able to run unaffected. Another example of hardware virtualization is "slicing": carving out precise portions of the system to run in a "walled garden," such as allocating a fixed 25% of CPU resources to bulk encryption. If there are no processes that need to crunch numbers on the CPU for block encryption, then that 25% of the CPU will go un-utilized. If too many processes need mathematical computations at once and require more than 25%, they will be queued and run as a FIFO buffer because the CPU isn't allowed to give out more than 25% of its resources to encryption. This type of hardware virtualization is sometimes referred to as pre-allocation. Asymmetric multiprocessing is a form of pre-allocation virtualization where certain tasks are only run on certain CPUs. In contrast, symmetric multiprocessing is a form of dynamic allocation, where CPUs are interchangeable and used as needed by any part of the management system. Each classification of hardware virtualization is unique and has value, depending on the implementation. Pre-allocation virtualization is perfect for very specific hardware tasks, such as offloading functions to a highly optimized, single-purpose chip. However, pre-allocation of commodity hardware can cause artificial resource shortages if the allocated chunk is underutilized. Dynamic allocation virtualization is a more standard approach and typically offers greater benefit when compared to pre-allocation. For true virtual service provisioning, dynamic resource allocation is important because it allows complete hardware management and control for resources as needed; virtual resources can be allocated as long as hardware resources are still available. The downside to dynamic allocation implementations is that they typically do not provide full control over the dynamicity, leading to processes which can consume all available resources .A another classification to them includes which most researches believe and first to come in mind are full virtualization and para virtualization.The terms snapshot ,Teleportation etc refer to hardware virtualization .Hypervisors play the crucial role in determining the type ,usefulness and working of hardware virtualization

F.)**Storage Virtualization**

The most important one or rather to say a crucial one is the storage virtualization. Day by day increase of large chunks of data lead to data centers ,cloud computing etc is not new to us .The general classification of them are in two categories: Block virtualization is best summed up by Storage Area Network (SAN) and Network Attached Storage (NAS) technologies: distributed storage networks that appear to be single physical devices. Under the hood, SAN devices themselves typically implement another form of Storage Virtualization: RAID. iSCSI is another very common and specific virtual implementation of block virtualization, allowing an operating system or application to map a virtual block device, such as a mounted drive, to a local network adapter (software or hardware) instead of a physical drive controller. The iSCSI network adapter translates block calls from the application to network packets the SAN understands and then back again, essentially providing a virtual hard drive .File virtualization moves the virtual layer up into the more human-consumable file and directory structure level. Most file virtualization technologies sit in front of storage networks and keep track of which files and directories reside on which storage devices, maintaining global mappings of file locations. [5]

IV. OVERVIEW OF SNAPSHOT

The software industry is evolving to highly modern and efficient service oriented architectures that are targeted to provide service to increasingly large numbers of end users. In order to provide appropriate scale-out any kind of hiccups in the service infrastructure needs to be prevented as well avoided to. Currently, one of the most common bottlenecks in these information systems is the huge chunks of database that stores the persistent data. This is major cause of the reasons why database replication has received a lot of attention in the last few years and Data science has got a huge boost. Basic approach to safely backing up live data is to temporarily disable write access to data during the backup, either by stopping the accessing applications or by using the locking API provided by the operating system to enforce exclusive read access. This is tolerable for low-availability systems . High-availability 24/7 systems, however, cannot bear service halts. To avoid downtime, high-availability systems may instead perform the backup on a snapshot—a read-only copy of the data set frozen at a point in time—and allow applications to continue writing to their data. In some systems once the initial snapshot is taken of a data set, subsequent snapshots copy the changed data only, and use a system of pointers to reference the initial snapshot. This method of pointer-based snapshots consumes less disk capacity than if the data set was repeatedly cloned. Read-write snapshots are sometimes called branching snapshots, because they

implicitly create diverging versions of their data. Aside from backups and data recovery, read-write snapshots are frequently used in virtualization , sandboxing and virtual hosting setups because of their usefulness in managing changes to large sets of files.[6]Snapshot are implemented in volume mangers ,file system, virtualization and databases. The implementation are in general .In recent years new patents and research have made new techniques and implementation of snapshot. Here the differential snapshot are i.e basically two types of differential snapshots: copy-on-write and redirect-on-write are explained as follows:

A)Copy-on-write snapshot: At the time when the snapshot is created, a small volume of space is allocated as a snapshot volume with respect to the source volume space. Upon the first write to a data block after the snapshot, the original data of the block is copied from the source volume to the snapshot volume. After copying, the write operation is performed on the block in the source volume. As a result, the data image at the time of the snapshot is preserved. The combination of the source volume and the snapshot volume presents the point-in-time image of the data. After the snapshot is created, all subsequent read I/Os are performed on the source volume.

B)**Redirect-on-write snapshot**:Copy-on-write requires three I/O operations upon the first write to a block (1) read the original block from the source volume, (2) write the original block to the snapshot volume, and (3) write the new data in the source volume. These I/O operations are done at production time, which has effects on the performance. To overcome this, one can do redirect-on-write that leaves the original block in the source volume intact and the new write operation is performed on the snapshot volume. This eliminates the extra I/O operations of the copy-on-write method. After the snapshot, all subsequent write I/Os are performed on the snapshot volume while read I/Os may be from source volume or snapshot volume depending on whether the block has been changed since the snapshot. The point-in-time image of the data at the time of a snapshot is the source volume itself since the source volume has been read-only since the snapshot time. The source volume will be updated at a later time, hopefully not in production time, by copying data from the snapshot volume.[7] In simple terms copy on write follows a snail based process first reading ,then making replica and then writing on original and finally updating while redirect on write does it in simple manner by doing the later operation on replica copy.

Now a brief introduction of two snapshot file technologies used in Linux operating system and Microscoft operating system. Shadow Copy(also known as Volume Snapshot Service,Volume Shadow Copy ServiceorVSS), is a technology included in Microsoft Windows that allows taking manual or automatic backup copies or snapshots of computer files or volumes, even when they are in use. It is implemented as a Windows service called theVolume Shadow Copyservice. Shadow Copy technology requires the file system to be NTFS an obivo reason is at it as Windows system implementation ,to be able to create and store shadow copies. Shadow Copies can be created on local and external (removable or network) volumes by any Windows component that uses this technology, such as when creating a scheduled Windows Backup or automatic System Restore point. LVM is a logical volume manager for the Linux kernel that manages disk drives and similar mass-storage devices.LVM is commonly used for the following purposes:Managing large hard disk farms by allowing disks to be added and replaced without downtime or service disruption, in combination with hot swapping.On small systems (like a desktop at home), instead of having to estimate at installation time how big a partition might need to be in the future, LVM allows file systems to be easily re sized later as needed. Performing consistent backups by taking snapshots of the logical volumes. Creating single logical volumes of multiple physical volumes or entire hard disks (somewhat similar to RAID 0, but more similar to JBOD), allowing for dynamic volume resizing.

If a comparison is done on both the technologies result would come as negative as both in present world of computing needs to be updated and most researchers believes the new version of LVM is better in comparison to its windows counterpart .Most of them are just hypothesis actual research and implementation of them is to be tested and verified

V. CONCLUSIONS

In this paper ,an overlook of the buzzwords virtualiation and snapshot and basics are explained. This paper gives us brief insight on a aura of information and research needed to be done on the field of snapshot and virtualization. Based on hypothesis and theory we firmly believe that virtualization is now basic alphabet in computing world. Seeing the increase of data day by the day the present technologies of snapshot need to be enhanced .In the present computing field the knowledge of these buzzwords are essential and the doors of imagination and implementation are open in the field of virtualization

VI. ACKNOWLEDGMENT

I would take this opportunity to thank Prof. Vatsal Shah for helping in the topic to understand the concepts and encouraging me to take this topic. Also I am thankful to all of my teaching and non-teaching staff in the department and my friends for extending their warm kindness and help. Finally I would like to thank my parents without their blessings

and support I would not have been able to accomplish my goal. Finally, I thank the almighty for giving the guidance and blessings.With all this assistance little for remains for which I can take credit.

REFERENCES

[1] PETR SUBA, JOSEF HORALEK, MARTIN HATAS, **"Comparison of Technologies for Software Virtualization"**,Recent Advances in Applied & Biomedical Informatics and Computational Engineering in Systems Applications, Pages 160-165 ,2011.

[2] Garimella, Neeta,**"Understanding and exploiting snapshot technology for data protection, Part 1: Snapshot technology overview",**2006.

[3] ALAN MURPHY, **" VIRTUALIZATION DEFINED - EIGHT DIFFERENT WAYS"**,F5 white paper.

[4] "**Optimized Storage Solution for Enterprise Scale Hyper-V Deployments**". Microsoft. March 2010.

[5] Weijun Xiao, Yinan Liu, and Qing (Ken) Yang, "**Implementation and Performance Evaluation of Two Snapshot Methods on iSCSI Target Storages**",Proc. of NASA/IEEE Conference on Mass Storage Systems and Technologies. 2006.

[6] Christine, Leja,, Chair Richard C, Barnier Charles ,L. Brown, Paul F. Dittmann, Paul Koziel, Mark Welle J.T. Westermeier, **"Virtualization and Its Benefits"** AITP – Research and Strategy Advisory Group**.**

[7]. Menascé, Daniel A. "**Virtualization: Concepts, applications, and performance modeling**."Int. CMG Conference. 2005.

[8] Morty Eisen ,"**Introduction to Virtualization**"The Long Island Chapter of the IEEE Circuits and Systems (CAS) Society,Marcum Technology

[9] Serrano, Damián, et al. "**Boosting database replication scalability through partial replication and 1-copy-snapshot-isolation.**"Dependable Computing, 2007. PRDC 2007. 13th Pacific Rim International Symposium on. IEEE, 2007.

[10] Kim, Chang-Soo, et al. "**A method for enhancing the snapshot performance in SAN volume manager.**"Advanced Communication Technology, 2004. The 6th International Conference on. Vol. 2. IEEE, 2004.

