International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

A Novel Hadoop Based Virtualization Automated Framework

G.J.Sunny Deol , O.Nagaraju

Abstract*: With the huge development in relocation to distributed computing for customary applications, the interest inHadoop VIRTUALIZATION is the new requirement for further research and improvements. The extreme interest for Hadoop VIRTUALIZATION in distributed computing is been tended to with the innovation improvement and regarding virtual machines. Significantly the equipment segments the executives of virtual machine is impressively achieving the pick of research with the ongoing headways by different organizations and open source explores like AlphaVM, Hyper-V, Integrity Virtual Machines, JPC (Virtual Machine), PowerVM, Sun xVM, VMware Workstation and z LPARs. Another most imperative segment of Hadoop VIRTUALIZATION and virtual machine picture stockpiling is been the bottleneck for the further headway. The impediments are being tended to by bunching and capacity territory organizing, anyway these current arrangements are no counterpart for the modern evaluation interest for the VM picture stockpiling necessities. Subsequently in this work we propose a cloud based virtual machine stockpiling structure intended for huge scale arrangement fulfilling alternate requirements for capacity highlights like replication and estimations of execution enhancements. This work additionally exhibits the headways in execution for Hadoop VIRTUALIZATION the board.

***Keywords* : Virtual Machine, Migration, Hadoop VIRTUALIZATION Bottlenecks, VMMS, VDMS, VM Image, Storage Architecture, Performance Comparison, Hadoop VIRTUALIZATION Management Framework**

I. INTRODUCTION

The centre of distributed computing for giving the versatile foundation is Hadoop VIRTUALIZATION. Hadoop VIRTUALIZATION is utilized to control the fundamental processing and correspondence foundation so as to make numerous occasions of similar assets to be dedicatedly apportions to different clients or client applications [1].

The Hadoop VIRTUALIZATION as a system enables the specialist co-ops to make diverse sorts of registering situations to coordinate the client prerequisites on an equivalent server. Hadoop VIRTUALIZATION enables the specialist organizations to address the issue for making the adaptable and application depended conditions.

Revised Manuscript Received on December 30, 2019.
***** Correspondence Author
G.J.Sunny Deol *, Research Scholar, Department of C.S.E, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India.
Dr.O.Nagaraju, Asst. Professor, Department of Computer Science, Govt Degree College, Macherla, Andhra Pradesh, India.

In the meantime, the client can likewise lessen the expense for keeping up processing and capacity related foundation on side. Most of the specialist co-ops convey the utilization of virtual machines so as to isolate real equipment from the registering situations. To comprehend the appropriateness of Hadoop VIRTUALIZATION through virtual machines, we consider the way that if there should arise an occurrence of private and open cloud the necessities for numerous frameworks, expansion of assets and diminishing the expense for usage is the prime vital factor. If there should be an occurrence of a private cloud the client will in general lease the foundation from a seller or now and again have the framework from possess premises.

If there should arise an occurrence of the on premises facilitating the equivalent virtual machine can be utilized for assortment of purposes and if there should arise an occurrence of off premises facilitating the supplier can utilize a similar design for numerous cu0stomers. Consequently,in both the cases, the usage of Hadoop VIRTUALIZATION through virtual machines lessens the expense [2].

In the opposite side for open cloud condition, the client can pick different arrangements for virtual machines to suite the necessities. In this manner, the cost execution for overseeing independent and custom design of equipment lessens for the suppliers.

So as to comprehend the upgrades required for Hadoop VIRTUALIZATION, we comprehend the equipment segments impacts the execution of Hadoop VIRTUALIZATION. The real parts for Hadoop VIRTUALIZATION are like the equipment assets required for nonexclusive processing[10]. The foundation manages the processor the executives, memory the executives, fringe the board, arrange the executives and capacity the board[11]. The capacity the executives for virtual machines assumes the significant job for execution improvements. The detail ponder is been led in the later piece of this work.

The capacity of virtual machine picture is the centre of this examination. Preceding the investigation of virtual machine picture positions, we comprehend the detail setups of virtual hard plate. The sending of virtual hard plate permits the specialist organizations, designers and clients to imitate numerous situations on single hard circle.

Additionally, the ongoing headway is inquiring about demonstrates the advantages of appropriateness of virtual hard circle as strong to transport on various host server, give depiction reinforcements and in time recuperation, give security lastly picture based administrations. The subtleties of the virtual machine designs are additionally been considered in this work.

We additionally consider the virtual machine screen programming and comprehend the extent of upgrade in the equivalent.

Virtual machine the board utilizing programming segments are famous and exhibited utilizing numerous inquiries about[15]. The virtual machine screen or VMM are the product segments dwelling on the host server to oversee and screen the virtual machines. The product gives a stage to the server based host working frameworks to have other processing situations for all intents and purposes. The detail investigation of the VMMs is additionally led in this work.

The ongoing investigations have exhibited that the capacity of virtual machine pictures is the greatest bottleneck of the progression and diminishing cost factors[16]. We have comprehended the ramifications of the arrangements regarding sending stockpiling zone systems. Anyway, the organization of SAN so as to deal with the capacity conveyances isn't practical.

Subsequently in this work, we attempt to propose a novel methodology for virtual machine appropriated picture stockpiling designs with effective administration and observing. The significant results of this work are an administration stage for overseeing virtual machine replication and capacity utilizing shared exchange convention and concentrate the upgrades in exhibitions [3] [4].

I. PERFORMANCE COMPARISON MATRIX

The execution of a Virtual Machine should be assessed dependent on different components to comprehend the nature of administration. Amid the administration level understanding marking process both the client and the specialist co-op consent to the recorded execution parameters. Henceforth observing the execution is one of the key parameter to screen the equivalent. Subsequently we comprehend and propose a lot of parameters to screen the equivalent. The parameters are characterized considering by and large prerequisites from the clients and perusing different administration level assertions [Table – 1].

TABLE I
VIRTUAL MACHINE PERFORMANCE EVALUATION MATRIX

Type	Name	Description
Dash Board Parameters (Over All Monitoring)	Name of VM	Unique Name of the VM
	Type	Host Architecture Type
	Overall Health Indicator	Running, Stopped, Critical
	Last Backup	Last Backup Date and Time
	Total Availability	Time of Total Availability
	Memory Utilization	Total Memory Utilization time
	Disk Utilization	Total Disk Utilization in GB
	Network Utilization	Total Network Utilization in time
Memory Parameters	Active Memory	Amount of Active Memory in GB
	Over heading Memory	Amount of Over heading Memory in GB
	Swappable Memory	Amount of Swappable Memory in GB / MB
	Total Shared Memory	Amount of Total Shared Memory
	Memory Temperature	Temperature of the Memory Units
Storage Parameters	Container Name	Unique name of the Storage Container
	Container Size	Container Size in GB
	Container Utilization	Container Utilization in GB
Networking Parameters	Network Card ID	Unique id for the Network Interface Card
	Up Time	Total Up Time
	Down Time	Total Down Time
	IP Address	Unique assigned IP Address
	MAC Address	Unique assigned MAC Address
	Data Transfer Rate	Data Transfer Rate in Megabytes per second
Peripheral Parameters (I/O Device)	Device ID	Unique Device ID
	Type	Read or Write
	Read Count	Number of Read Operations
	Write Count	Number of Write Operations

After defining the novel matrix for virtual machine performance evaluation, we test the performance of the proposed system, which is elaborated in further section of this same work.

II. VIRTUALIZATIONUSING HADOOP STACK

In this section of the work, the possibilities of virtualization based on the Hadoop stack is analysed.

A. Hadoop VIRTUALIZATION using Hardware:

Hadoop VIRTUALIZATION utilizing Virtual Machines empowers to segregate the physical equipment segments from the product stack and builds the profitability talked about in early part on this work. During the time spent Hadoop VIRTUALIZATION, we comprehend the standard structure lives on the physical equipment and the virtual machine repeats the equipment legitimately. In the virtual machine the working framework gets introduce and runs the applications [Figure – 1].

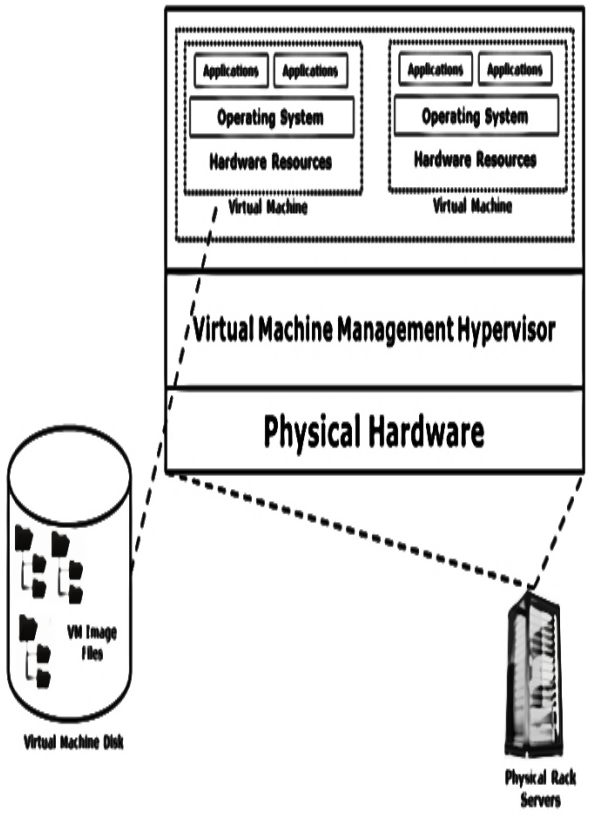

Figure. 1 Hardware Components for Hadoop VIRTUALIZATION

Here we comprehend the arrangements given by various organizations and by the autonomous pursuits, a base degree is accessible to go for the improvements. In this manner, we centre for the enhancements conceivable on the off chance that Hypervisor engineering[17].

B. Hadoop VIRTUALIZATION using Software Stack:

So as to give the improvements to Hypervisor innovation, here we comprehend the engineering. The Hypervisor is mostly in charge of controlling the virtual machines alongside dealing with the systems administration segments, record frameworks, virtual machine pictures, programming stack, the replication and other programming controller.

The Hypervisor is isolated into three noteworthy layers as control layer, organize layer and interfacing layer among physical and intelligent systems administration layers [Figure – 2].

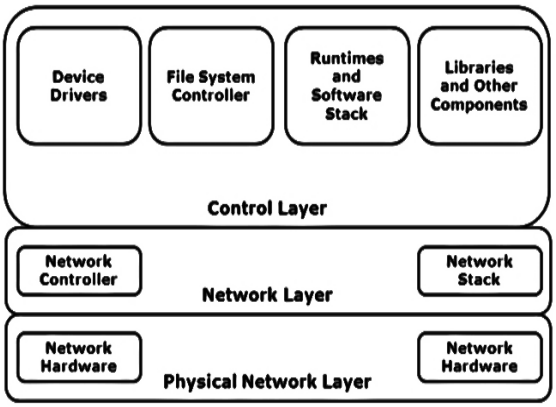

Figure. 2 Generic Architecture for Hypervisor

Subsequently we propose the novel methodology for capacity and processing Hadoop VIRTUALIZATION the board. The detail of the proposed system is talked about in the later piece of this work [7 - 12].

III. PROPOSED COMPUTING HADOOP VIRTUALIZATION MANAGEMENT FRAMEWORK

After the detail comprehension of the virtual machine system and executions, we understand the requirement for development in Hypervisor structure. Significantly we comprehend the accompanying issues are by and large experienced in the clear majority of the well-known Hypervisor programming[18].

The circumstance where numerous physical equipment servers are used for same or different customers, at that point the observing and overseeing singular equipment is almost incomprehensible from a solitary firmware or Hypervisors as the Hypervisors are limited to explicit equipment sellers.

Henceforth we propose the Storage and Computing Virtualization Management Framework*[Figure – 3]*.

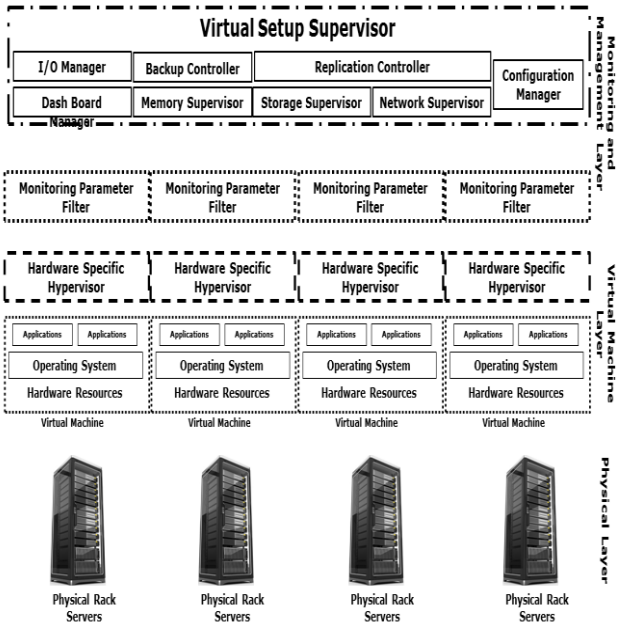

Figure. 3 Proposed Automated Framework

IV. RESULTS

The tale checking and the executive's system for the register and capacity Hadoop VIRTUALIZATION is been connected and the outcomes watched are tasteful. The system is being investigated for execution dependent on the parameters characterized on the past piece of this work.

The result of the finding is observed here [Table – 2]:

TABLE II
PROPERTY AVAILABILITY MATRIX

Parameter Observation	Availability of the Properties		
	Test Process – 1 (Duration – 60 Mins)	Test Process – 2 (Duration – 90 Mins)	Test Process – 3 (Duration – 200 Mins)
Name of VM	Available	Available	Available
Type	Available	Available	Available
Overall Health Indicator	Available	Available	Available
Last Backup	Available	No Continuous Availability	Available
Total Availability	Available	Available	No Continuous Availability
Memory Utilization	Available	Available	Available
Disk Utilization	Available	Available	Available
Network Utilization	Available	No Continuous Availability	No Continuous Availability
Active Memory	Available	No Continuous Availability	No Continuous Availability
Over heading Memory	Available	Available	No Continuous Availability
Swappable Memory	Available	Available	Available
Total Shared Memory	Available	No Continuous Availability	No Continuous Availability
Memory Temperature	Available	Available	Available
Container Name	Available	Available	Available
Container Size	Available	Available	Available
Container Utilization	Available	Available	Available
Network Card ID	Available	Available	Available
Up Time	Available	Available	No Continuous Availability
Down Time	Available	Available	No Continuous Availability
IP Address	Available	Available	Available
MAC Address	Available	Available	Available
Data Transfer Rate	Available	Available	Available
Device ID	Available	Available	Available
Type	Available	Available	Available
Read Count	Available	Available	No Continuous Availability
Write Count	Available	Available	No Continuous Availability

V. CONCLUSION

In this work,we considered the bottlenecks of Hadoop VIRTUALIZATION and in Hadoop VIRTUALIZATION adaptation as Access Delay, Acceptability by the researchers, Provisioning Deficiency and Interactivity Limitations. We further studied the components playing major role in Hadoop VIRTUALIZATION to be taken under consideration for further research as Hardware Components for Hadoop VIRTUALIZATION and Software Components or Hypervisor for Hadoop VIRTUALIZATION and identified that the dependencies and deficiencies in hypervisors. This work also considers the virtual machine image storage format in order to achieve better understanding of nature of the virtual disk files as Fixed Length Image File Format, Dynamic Length Image File, Dependent Image File and Linked Image File and realized the most suitable image formats.

REFERENCES

1. J. Rao, Y. Wei, J. Gong and C.-Z. Xu "Qos guarantees and service differentiation for dynamic cloud applications", IEEE Trans. Netw. Serv. Manag., vol. 10, no. 1, pp.43 -55 2013

2. Li, A. Raghunathan and N.K. Jha, &ldquo,Secure Virtual Machine Execution under an Untrusted Management OS,&rdquo, Proc. Int',l Conf. Cloud Computing, pp. 172-180, July 2010.
3. Intel VT-D, http://www.intel.com/technology/Hadoop VIRTUALIZATION/technology.htm, 2012
4. BYTEmark, http://www.tux.org/mayer/linux/byte/bdoc.pdf, 2012
5. M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica and M. Zaharia, &ldquo,Above the Clouds: A Berkeley View of Cloud Computing,&rdquo, Technical Report UCB/EECS-2009-28, http://www.eecs. berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html, 2010
6. T. Deshane, Z. Shepherd, J.N. Matthews, M. Ben-Yehuda, A. Shah and B. Rao, &ldquo,Quantitative Comparison of Xen and KVM,&rdquo, Xen Summit Boston 2008, http://xen.org/xensummit/xensummit_ summer_2008.html, 2010
7. Y. Dong, X. Yang, X. Li, J. Li, K. Tian and H. Guan, &ldquo,High Performance Network Hadoop VIRTUALIZATION with SR-IOV,&rdquo, Proc. IEEE 16th Int',lSymp. High Performance Computer Architecture (HPCA), pp. 1-10, 2010
8. P. Padala Automated management of virtualized data centers, 2010 :Univ. of Michigan
9. T. Patikirikorala, A. Colman, J. Han and L. Wang "A systematic survey on the design of self-adaptive software systems using control engineering approaches", Proc. Symp. Softw. Eng. Adaptive Self-Manag. Syst., pp.33 -42 2012
10. Lakshman Narayana Vejendla and Bharathi C R,(2017),"Using customized Active Resource Routing and Tenable Association using Licentious Method Algorithm for secured mobile ad hoc network Management", Advances in Modeling and Analysis B, Vol.60, Issue.1, pp.270-282.
11. Bikku, T., Rao, N. S., &Akepogu, A. R. (2016). Hadoop based feature selection and decision making models on big data. Indian J. Sci. Technol, 9(10).
12. X. Wang and Y. Wang "Coordinating power control and performance management for virtualized server clusters", IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 2, pp.245 -259 2011
13. X. Wang, M. Chen and X. Fu "MIMO power control for high-density servers in an enclosure", IEEE Trans. Parallel Distrib. Syst., vol. 21, no. 10, pp.1412 -1426 2010
14. T. Patikirikorala, A. Colman, J. Han and L. Wang "A multi-model framework to implement self-managing control systems for QoS management", Proc. Int. Symp. Softw. Eng. Adaptive Self-Manag. Syst., pp.218 -227 2011.
15. Lakshman Narayana Vejendla , A Peda Gopi and N.Ashok Kumar,(2018)," Different techniques for hiding the text information using text steganography techniques: A survey", Ingénierie des Systèmes d'Information, Vol.23, Issue.6,pp.115-125.
16. Bikku, T., Nandam, S. R., &Akepogu, A. R. (2017). An Iterative Hadoop-Based Ensemble Data Classification Model on Distributed Medical Databases. In Proceedings of the First International Conference on Computational Intelligence and Informatics (pp. 341-351). Springer, Singapore.
17. A Peda Gopi and Lakshman Narayana Vejendla, (2019)," Certified Node Frequency in Social Network Using Parallel Diffusion Methods", Ingénierie des Systèmes d' Information, Vol. 24, No. 1, 2019,pp.113-117.
18. Lakshman Narayana Vejendla and A Peda Gopi, (2019)," Avoiding Interoperability and Delay in Healthcare Monitoring System Using Block Chain Technology", Revue d'Intelligence Artificielle , Vol. 33, No. 1, 2019,pp.45-48.

AUTHORS PROFILE

G.J.Sunny Deol a Research Scholar of Computer Science & Engineering in Acharya Nagrajuna University, Nagarjuna nagar. He is having 10 years of teaching and research experience He completed his B.tech and M.tech degree under JNTU & JNTUK.He is engaged in **"A Novel Hadoop Based Virtualization Automated Framework** " Project developing process**.**

Prof. O.Nagaraju received Ph.D(CSE) and M.E (C.S.E) from Acharya Nagrajuna University , Nagarjuna nagar. He is having 13 years of Teaching Experience in various Engineering Colleges. At present he is working as Asst.Professor, Department of computer science, Government Degree College,Macherla-522426,