USING VIRTUALIZATION FOR DISTRIBUTED COMPUTING

[1]ANU KAUL, [2]MEENA GUPTA, [3]KHUSHWANT KAUR

[1,2,3]Department of Computer Science, Sri Guru Gobind Singh College, Sector 26, Chandigarh
E-mail: [1]kaul7anu@gmail.com, [2]meena.gupta355@gmail.com, [3]kkhushwantkaur@yahoo.com

Abstract— The past few decades have seen many changing trends in distributed computing systems in the form of grid, cloud, utility and cluster. With the advent in technology, there has been an increase in the demand for deployment of a robust distributed network for maximizing the performance of such systems and minimizing the infrastructural cost. In this paper we have discussed various levels at which virtualization can be implemented for distributed computing which can contribute to the increased efficiency and performance of distributed computing.

Index Terms— Virtualization, Distributed computing, Virtual machine, Virtual Machine Monitor, Utility Computing, Computational Grids, Transactional Grids, Data Grids, Storage Grids.

I. INTRODUCTION

The technological trends have always been aiming towards development of such systems that can run many processes simultaneously, can share the available resources across various users, groups and organizations and that can attain maximum performance from the available infrastructure. Distributed computing is a complex process when comes to its development and deployment because of its transparency in sharing resources and parallel execution of multiple processes.

Most data centers today have a three- or four-tier hierarchical networking structure. Three-tier networking architectures were designed around client-server applications and single-purpose application servers. This type of network architecture, however, is becoming problematic for the data center. Now-a-days application environments are more distributed, often with multiple tiers, and oriented toward service delivery [1]. Main concern of distributed computing is to improve energy efficiency and resource utilization. The technological improvements in the form of distributed computing, grid computing and cloud computing in WAN has made it possible to aggregate distributed resources.

For improvising the system's maintenance and management, rapid advancement in system virtualization at various levels is taking place. There is need to employ a performance enhancer mechanism to maximize the reliability, scalability and fault tolerance of the distributed systems.

II. DISTRIBUTED COMPUTING

Distributed computing refers to more than two computers interacting to share the processing of a single program to implement resource management, performance enhancement, increased reliability, scalability, heterogeneity, transparency, concurrency, security and privacy.

Distributed computing is a method of computer processing in which different parts of a program are run simultaneously on two or more computers that are communicating with each other over a network. Distributed computing is a type of segmented or parallel computing, but the latter term is most commonly used to refer to processing in which different parts of a program run simultaneously on two or more processors that are part of the same computer. While both types of processing require that a program be segmented—divided into sections that can run simultaneously, distributed computing also requires that the division of the program take into account the different environments on which the different sections of the program will be running [2].

There are several definitions for distributed systems. Coulouris defines a distributed system as "a system in which hardware or software components located at networked computers communicate and coordinate their actions only by message passing" [3]; and Tanenbaum defines it as "A collection of independent computers that appear to the users of the system as a single computer" [4]. Leslie Lamport said that "A distributed system is one on which I cannot get any work done because some machine I have never heard of has crashed" [5].

III. VIRTUALIZATION

Virtualization is the most effective way to reduce IT expenses and boost efficiency large to mid size and small organizations. Virtualization lets you run multiple operating systems and applications on a single server, consolidate hardware to get vastly higher productivity from fewer servers and simplify the management, maintenance, and the deployment of new applications.

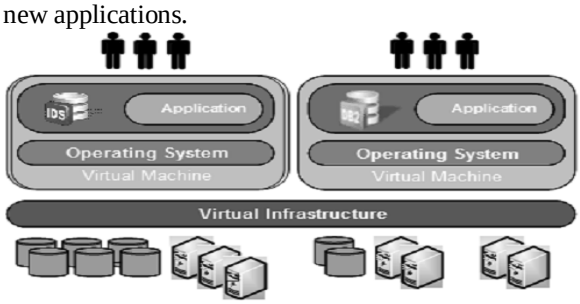

Fig1: Concept of Virtualization

Virtualization creates a virtual version of a device or resource, such as a server, storage device, network or even an operating system where the framework divides the resource into one or more execution environments. The term virtualization is now associated with a number of computing technologies including the following [6]:

- *Storage Virtualization*: the amalgamation of multiple network storage devices into what appears to be a single storage unit.
- *Server Virtualization*: the partitioning a physical server into smaller virtual servers.
- *Operating System Virtualization*: a type of server virtualization technology which works at the operating system (kernel) layer.
- *Network Virtualization*: using network resources through a logical segmentation of a single physical network.
- *Application Virtualization*: each application brings down its own set of configurations on-demand, and executes in a way so that it sees only its own settings. This leaves the host operating system and existing settings unaltered.

Virtualization creates an environment wherein it emulates and imitates various hardware components like CPU, OS, software, I/O devices and storage devices to numerous virtual machines (VM). Each node in the distributed system is a virtual machine running independently. The virtual machine monitor (VMM) administers the privacy of all the connected VMs so that each VM can perform the task assigned to it without getting affected by another VM.

IV. LEVELS FOR IMPLEMENTING VIRTUALIZATION FOR DISTRIBUTED COMPUTING

Distributed virtualization means transparent sharing of resources by many clients. The clients are neither aware of each other nor of various available resources. A larger composite environment can be created by assimilating smaller systems.

There are various levels at which virtualization can be implemented to increase the performance of a distributed network. Each level provides a unique direction to create a more reliable, robust, cost effective and efficient distributed system. The distributed computing can implement the technique of virtualization on one or more levels to maximize the performance. Each level has some issues and concerns which are also discussed along.

A. Application Virtualization [7]:

Various organizations share the access of applications by as remote services. It leads to reduction in cost for software and infrastructure. The main concern for this level is to maintain the security across multiple users.

B. Utility Computing [7] :

Resources from various connected data centers are made available to be used as and when necessary. These data centers are also based on distributed computing so that infrastructural cost can be minimized. Virtualization can allocate resources efficiently among various clients. Major concern on this level is the proper implementation of resource management.

C. Grids [7] :

- *Computational Grids*: Transparent sharing of computational server resources among multiple groups across or within an organization. Cross-organization management has to be carefully implemented.
- *Transactional Grids*: To support high performance transactional applications, the distributed hardware and software resources can be shared. Its implementation lacks standardization.
- *Data Grids*: Data servers can be shared between various groups for easier access to distributed data. Virtualization at this level will help in managing and maintaining these groups in a more sophisticated manner. Main issue is the maintenance and consistency of the metadata.

D. Virtual servers [7]:

Virtual Server can use multiple operating systems and applications giving it functionality of multiple servers into one server. Such server in the networks increases the fault tolerance and makes it energy and cost efficient.

E. Virtual machines [7]:

These are the nodes running their own CPU and OS in an independent environment for execution of various applications on portable platform.

F. Storage grids and utilities [7]:

Storage virtualization creates one virtual environment of multiple storage devices which makes back up, migrating data and storage expansion for multiple clients very efficient.

CONCLUSION

Virtualization in distributed computing brings flexibility, scalability, portability and cost advantage to an organization. The coupling of virtualization and distributed computing has opened many options for the large to small sized organizations by maximizing the performance by utilizing the resources efficiently thereby resulting into minimizing the infrastructural cost.

There are some issues, challenges and concern which are to be considered and ways must be explored to overcome the serious issues like security, privacy, network management, data consistency, user and group management, maintenance and functionality of data when access is shared among different groups and organizations.

This approach can connect wide variety of computer types and wide range of computing resources to create vast "virtual" reservoirs of computers which can be shared among multiple users globally. The nature of this paper is exploratory. Various levels of distributed computing have been discussed at which virtualization can be implemented to harness unused computing power in the computers being used in an organization.

REFERENCES

[1] Hp Technology Brief, Effects of Virtualization and Cloud Computing on Data Center Networks.

[2] Attiya, Hagit and Welch, Jennifer (2004). Distributed Computing: Fundamentals, Simulations, and Advanced Topics. Wiley-Interscience. ISBN 0471453242.

[3] G. Couloris, J. Dollimore, and T. Kinberg, Distributed Systems – Concepts and Design, 4th Edition, Addison-Wesley, Pearson Education, UK, 2001.

[4] A. Tanenbaum and M. Van Steen, Distributed Systems: Principles and Paradigms, Prentice Hall, Pearson Education, USA, 2002.

[5] Krishna Nadiminti, Marcos Dias de Assunção, and Rajkumar Buyya, Distributed Systems and Recent Innovations: Challenges and Benefits, The University of Melbourne, Australia

[6] http://www.webopedia.com/TERM/V/virtualization.html

[7] Bob Marcus, Distributed Virtualization for Net-Centric Operations, link: home.comcast.**net**/~**bobmarcus**1/dv-draft.**pdf.**

★★★