Design of vFarm and it's Reporting

Dike Siddhesh Sakharam[1] and Radhika D.Joshi[2]

1Dept. of Electronics and telecommunication College of Engineering Pune
[2]Dept. of Electronics and telecommunication College of Engineering Pune
E-mail: [1]dikesiddheshcoep14@gmail.com, [2]rdj.extc@coep.ac.in

Abstract—*Cloud computing has mainly three features, Infrastructure as a Service (IAAS), Software As A Service (SAAS) and Platform As A Service (PAAS). IAAS is used to access resources from virtual infrastructure. IAAS is designed based on the concept of Virtualization. With the help of virtualization technology, available physical resources can be used effectively over virtual environment. For providing services through IAAS, a virtual environment is developed and that virtual environment is called 'virtual farm'. After design of virtual farm (vFarm) different parameters are checked such as latency, throughput and CPU utilization. The designed virtual farm can be managed centrally by an administrator who may be technical or non-technical person. The utilization of vFarm parameters should be monitored and reported effectively so that managing vFarm gets easier. This paper emphasizes more on design of vFarm and reporting tool.*

Keywords: *Virtualization,datacenter,cluster,host,virtual machine, CPU utilization,management,virtual farm.*

1. INTRODUCTION

The idea of virtualization can be used for most of the IT components such as networks, storages, servers, software applications and operating systems. Among all above virtualization techniques , server virtualization is the liked and demanding model in IT industries. Virtualization is the most significant for security and infrastructure management [1].Server virtualization implements one virtual layer on top of a physical system so that all physical devices are not visible to clients and generate a pooling environment for creating several numbers of independent virtual machines. So we can say that server virtualization is the method of representing logical servers of computing resources so that resources can be accessed with more advantages over the original configuration. Server virtualization is used for creation of multiple independent operating environment (logical or virtual servers) to run on physical server i.e. dividing one physical server components into multiple virtual machine components [2]. A virtual server farm is a networking environment that consist multiple applications and infrastructure servers running on large number of physical servers using a server virtualization program such as xen server, KVM, Microsoft Virtual Server or VMware.

2. LITERATURE SURVEY

Methods discussing usage of virtual cluster are illustrated in the literature. Virtual cluster is designed for teaching laboratory and performance are checked on a cluster in paper [3]. In this paper cluster performance is compared based on single trunked NIC (Network Interfacing Card) and dual NIC card. In paper [4], virtual lab is designed for network administration and for educational needs. This paper explains about designed structure of lab and used of designed lab in education sector. Pooja and Asmita Pandey explain performance of virtual machine based on SLA (System Level Agreement) violation and power consumption.Performance of virtual cluster, virtual lab and virtual machine to be measured, so that effective utilization of resources is carried out. In this paper, number of clusters and data centers are designed and assigning a network, storage pool to all datacenter is done. After designed of vFarm, for resource management and performance measurement, reporting framework is designed.

Difference between Virtual and physical server farm

In the physical server farm, every different function or application requires a different server. Physical server farm consists of large number of racks having a number of physical servers. These servers have pre-installed operating system which can run applications. Operating system layer in the virtual server is replaced by a hypervisor. The hypervisor allows to run multiple virtual machines on single hardware. Each virtual machine has its own operating system and applications running on it. Connecting all the virtual hosts together, a virtual datacenter is created. Large number of virtual datacenters combine together to form a virtual farm. Each virtual machine needs hardware resources like RAM, power, CPU etc. All resources are made available by hypervisor. There are going to be a multiple points of administration for physical farm and virtual farm. In physical farm, we have to access console of server, console of network devices like switches and routers and console of storage area network, individually. A virtual farm administration is easier than that of a physical server. Virtual farm is completely dependent on physical infrastructure. The strength of virtual farm depends on how strong the physical datacenter is.

3.   PROPOSED WORK

3.1 Implementation of virtual farm(Backbone)

Implementation of virtual farm involves capacity analysis, device selection, configuration and implementation, and testing and allocating for production. Configuration and implementation consists of different steps such as virtual farm configuration, network configuration and storage configuration.

3.2 Capacity Management

The capacity analysis depends on number of users that access the virtual farm and the number of virtual machines accessed by each user. There are two steps present in capacity analysis,

3.2.1 Bucketization

In this step, total number of virtual machines working under virtual farm is checked. Then collect the CPU capacity of individual virtual machine and decide memory, storage and network requirement for each virtual machine. Capacity of each virtual machine is dependent on the total number of running applications on the machine. Hence, essential to find total peak CPU utilization required to run the applications on an individual machine. Then after collecting data of all machines from virtual farm, the total CPU utilization can be calculated, required for virtual farm. Similar to CPU utilization, total RAM utilization, storage requirement and the number of network ports required in virtual farm can be estimated.
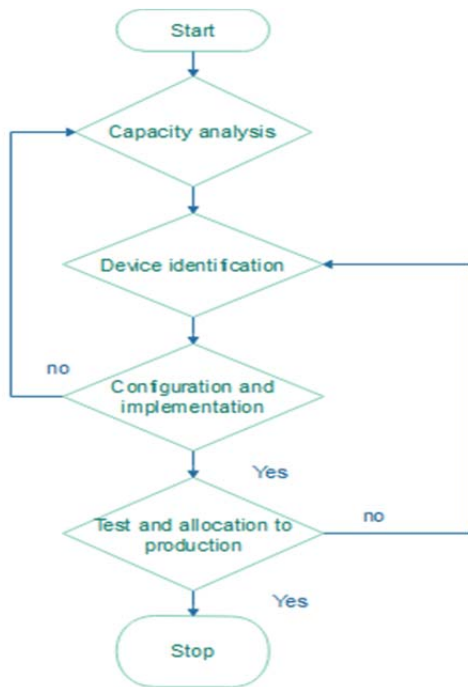

Fig. 1: Algorithm to design backbone of vFarm.

3.2.2 Tiering

In tiering, assigning different kind of application to different type of storage media is done. Different applications requires different memory, different CPU utilization.

Table 1: Tier configuration

Tier Level	Configuration
Tier 1	<=4 GB Mem, <=2vCPU, <=60GB HD
Tier 2	>4 - 8GB Mem, upto 4 vCPU, upto 120 GB HD
Tier 3	>8 - 12GB Mem, upto 6 vCPU, upto 200 GB HD
Tier 4	>12 - 16GB Mem, upto 8 vCPU, >200 GB HD

For instance, server requiring memory less than 4GB, virtual processor less than or equal to 2vCPU and hard disk capacity of 60GB is assigned to tier 1. Similarly, servers are allocated to tier2, tier3, tier4 based on the corresponding configuration details listed in Table (1).

3.2.3 Device Identification

The next step after tiering is device identification. The basic unit of virtual farm is a server. Along with the server, different devices such as switch, storage etc. is required for designing the virtual farm. Server selection depends on the total CPU utilization, total RAM and number of network ports. Depending on cost and specifications, servers are chosen. Servers have to interconnect for different operations such as fault tolerance and virtual motion. So switches are required to interconnect the servers. Layer 2 switches are mainly used to interconnect servers and layer 3 switches are used to connect different networks. To interconnect servers, VENA (Virtual Enterprise Network Architecture) enabled switches with iSCSI (Internet Small Computer System Interface) protocol are used. External storage is also used to enhance requirement of storage devices. External storage devices are accessible to the server. Storage is managed centrally. To manage storage devices fiber channel or iSCSI can be used. iSCSI is mostly used, since it is simple and easy to install.

3.2.4 Configuration And Implementation

Pre-implemented backbone network or separate backbone network can be cast off for designing a virtual farm.

Design of virtual farm has three different parts,

3.2.5 Server farm

Server farm has a number of clusters and each cluster consists of number of servers. All these servers are connected to each other. The server farm supports higher computing application than single physical machine.

3.2.6 Storage Farm

Storage farm consists of interconnection of number of external physical storage devices. These storage devices can be used as shared storage devices, which can be used among different virtual machines. Storage area network technology is used to connect storage devices.

3.2.7 Network

Network helps to connect physical devices in the infrastructure. Network of virtual farm is divided into different parts namely Management network, Virtual machine traffic network, virtual motion network and storage network.

3.2.8 Management network

Management traffic enables the vmkernel NIC for high availability. It helps in managing the configuration settings for each and every virtual device. vCenter management is done with management network.

3.2.9 V-motion traffic

V-motion allows maintenance of physical resources without any predefined indication. Virtual machines are moved from high resource utilizing server to low resource utilizing server. Also, v-Motion helps virtual machine to operate in case of server failures by porting the virtual machines to a running server.

3.2.10 Virtual machine traffic

Two different virtual machines from different host can communicate or exchange information called as virtual machine traffic. Not only communication between two virtual machines but communication between virtual machine and physical host is also possible.

3.2.11 Storage traffic

This is the high data rate traffic generated by storage input and output applications tool. Migration of storage resources for different virtual machines is also easy because of single storage network.
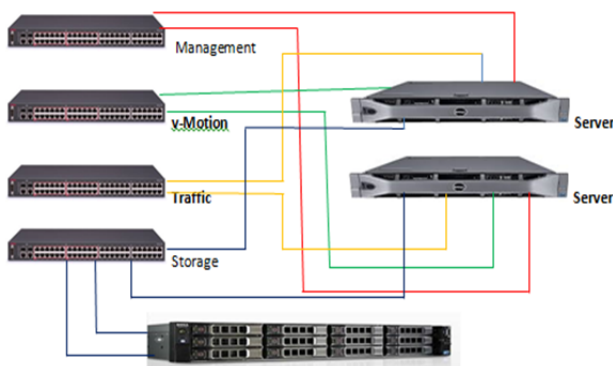

Fig. 3: Interconnection of physical components

3.3 Testing and allocating of vfarm resources to community:

The last steps towards Lab virtualization are testing infrastructure, modification in infrastructure if required and finally allocation of the infrastructure to users. While designing virtual farm, both management and traffic switches are used for transmission purposes so that, if one switch fails then another switch can perform the same function or not is checked. Different servers are allocated to different teams for various applications.

3.3.1 Implementation of virtual farm on backbone

Once the basic physical infrastructure has been setup and the connectivity tested, the next step is to install the management server component. The management server, also known as virtual center which is used for all management task.

Software used for management task,
- Vmware vCenter

- ESXi

ESXi is a virtualization server; it is a type 1 hypervisor. All the virtual machine or guest OS are installed on ESXi server. VMWare vSphere client allows administrators to connect to ESXi server and access or manage virtual machines. vSphere client is installed on client machine.

The use of hierarchical structure provides simplicity and ease in management. Hierarchical structure consists of datacenter, cluster, host and virtual machines. Host is the most important part of virtualization. Firstly, all host having ESXi installed are to be connected under the vCenter. The basic elements of clusters are network, memory (RAM), and storage elements. Cluster consists of large number of hosts. The cluster sizing is required to reach desired utilization across all elements of infrastructure. For deciding proper cluster size, virtual machine per datastore, number of EXSi connected to single datastore and a single datastore connected to multiple hosts must be considered.VMWare virtual center is used to manage virtual networking in virtual farm. On the same ESXi host, a virtual switch allows a large number of virtual machines to communicate with same protocol standards which are used in physical switch. Virtual Ethernet adaptor, which is used by each virtual machine and virtual switch are the key components of networking. There are total five types of virtual adaptors; anyone of them is used depending on the requirement. Virtual switch is used to connect virtual machine to external world. Virtual switch is connected to physical adaptor.
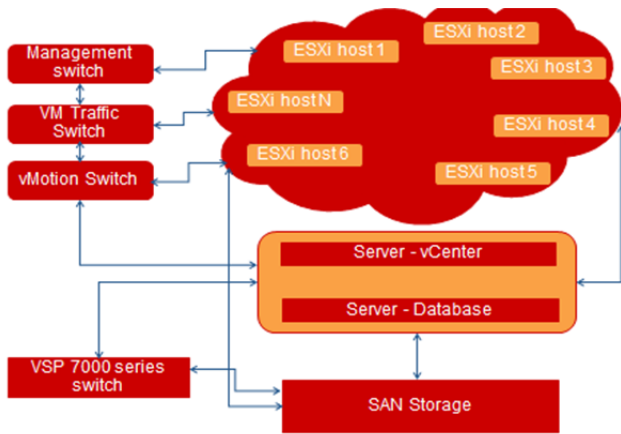
Fig. 4: Designed Virtual Farm

Different virtual switches are designed for vMotion, traffic, management and storage. One virtual switch is used to connect different virtual machines created under one host; the virtual machines can connect to each other if required.

3.3.2 Reporting

Creating a report consist of mainly four steps

3.3.3 Deciding an audience:

Audience means the person who is going to receive report of virtual farm. Audience may be an engineer or manager or executive. In this reporting, audience is manager. Managers require the detailed the information of reports effectively. Executives require the information regarding the steps to manage the virtual farm.

3.3.4 Deciding a final output:

In this step selecting the output format such as pdf, HTML or excel etc. is decided. In this case, pdf format is chosen. In pdf format, utilization of resources such as CPU, RAM, network and storage are represented in form of graphs, so that technical as well as non-technical person should understand clearly.

3.3.5 Interpretation of data:

In this step, acquiring data from vCenter server, filtering as well as analysis of data takes place. After analyzing data, interpretations are done for power management and resource management in the cluster of vFarm. In report, violating condition can be mentioned for effective management of vFarm.
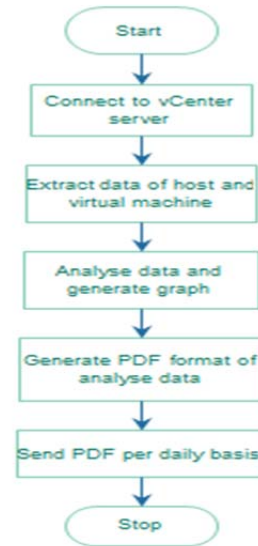

Fig. 5: Flow chart to design report

3.3.6 Automation:

In this step, scripts are written to automate reporting tool on daily basis and depending on violation of specific condition. After generating the report in pdf format, we have to send that report to appropriate manager or executive to make conclusion and perform specific task.

In report, utilization of CPU resources and memory resources of virtual farm on daily, weekly and monthly base is generated. Hence overall behavior resources utilization will understand and on that basis management is done. To manage power utilization in virtual farm, cluster wise power management is done. One cluster means one team in industry so team wise power management is done.So to manage power utilization algorithm is implemented in report which indicate effective scenario of cluster so that cluster may have less power consumption.

4. RESULTS

4.1 Latency Effect

In the above designed virtual farm, latency is plotted for different scenarios, such as PC to PC communication, PC to server communication, virtual to another host, virtual machine to another host virtual machine, virtual machine to same host and virtual machines between same hosts. Latency is a depending on processing power of host machine and number of devices present between communicating devices. As the packet size increases, latency increases. For communication from virtual machine to physical machine requires more latency than physical to physical communication. Communication between two virtual machines in a same host require minimum latency.
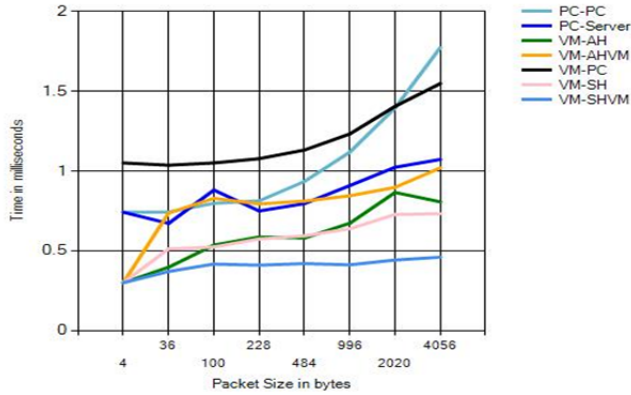
Fig. 6: Latency for different Scenario

4.2 CPU Utilization

CPU utilization is one of the most important factors. In this experiment, percentage of CPU utilization is checked with respect to number of virtual machines. So as the number of virtual machines increase, CPU utilization also increases linearly. It is checked for a UDP packet with a fixed length. Due adding one virtual machine in a host, near about 2 percent of CPU utilization increases for transmission of packets and as compared to transmitter, the receiver CPU utilization is low.
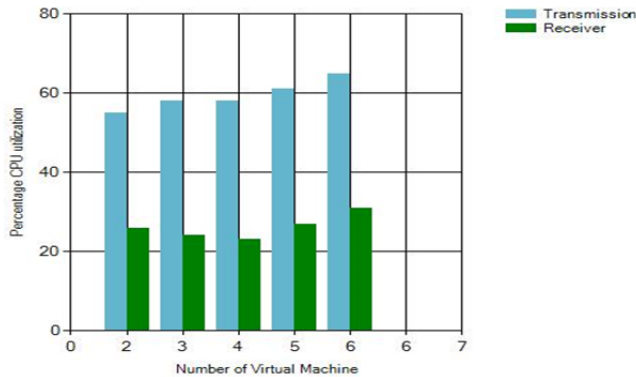

Fig. 7: CPU utilization for virtual machines

4.3 Throughput

In a virtual farm, throughput is an important factor. The effect on throughput is checked, with respect to varied number of virtual machines on a single host of cluster. It is observed that, as number of virtual machines increases, throughput decreases. If a single virtual machine is increased, then the throughput is decreased linearly with a specific factor and that factor is about 5 MBPS. In order to maintain throughput, host to virtual machines ratio is maintained. The ratio is dependent on CPU capacity in server host.
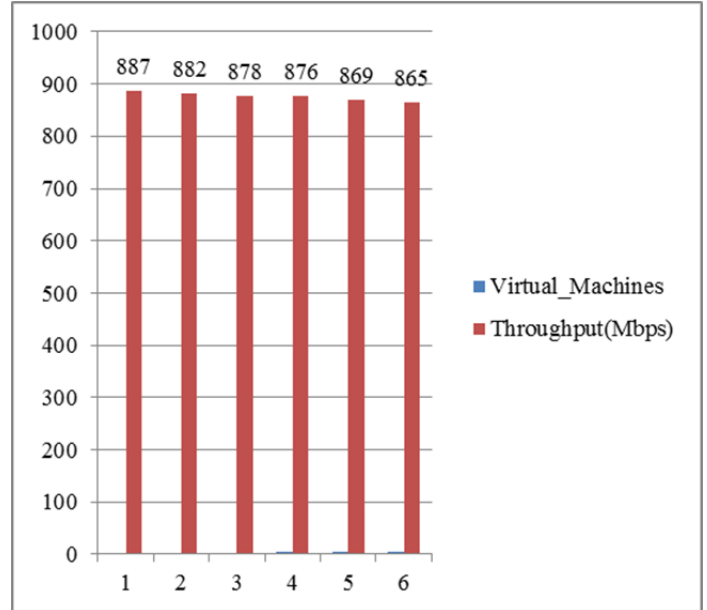

Fig.8. Throughput against number of VM

Then the throughput is checked, with respect to workload. Three workloads are used, named as bulk, file and default workload. In bulk, buffer size is 65536 bytes and send size is 8192 byte. In file buffer size is 8192 bytes and send size is 4096 byte. In default both sizes are same and value is 8192 byte. Bulk load has maximum throughput as compared to file and default workload. File and default workloads have near about the same throughput. Throughput is also checked for different condition such as virtual machine to personal computer, personal computer to virtual machine and personal computer to server. In above listed conditions, server to PC achieves a maximum throughput and PC to Virtual machine achieves minimum throughput.
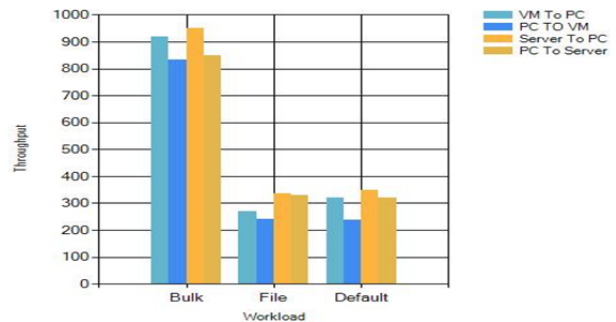

Fig. 9: Throughput against different workload.

Virtualization has some advantages and some disadvantages with physical devices. Advantages are latency and throughput while the disadvantage is virtualization lags in CPU utilization. More CPU power is required because one extra layer is introduced so processing at that layer, requires extra energy.

5. CONCLUSION

Efforts for designing vFarm effectively managing resources in an organization are illustrated. Power is effectively managed in vFarm with help of vMotion network. Efficient utilization of resources in a team is observed and report based on utilization is send to respective manager. The latency and throughput can be managed by keeping virtual machines to same host. Virtualization concept can be used in the new technologies such as Internet of things (IOT), cloud etc. For provisioning point of view, in IOT devices, virtual layer is most important. If one device fails then another virtual machine provides same functionality with in few second and that can be managed by virtualized layer. Virtualization can be used in IOT to reduce impact on memory, device access, latency and performance. In cloud to designed infrastructure and provide infrastructure as a service virtualization is used. To manage resources effectively virtualization is used as a new technology.

REFERENCES

[1] Mohammad Reza Ahmadi and Davood Maleki, "Performance Evaluation of Server Virtualization in Data Center Applications", International Symposium on Telecommunications, 2010, pp. 638 - 644.

[2] Eric Johnson, Patrick Garrity, Timothy Yates and Richard A. Brown, "Performance of a Virtual Cluster in a General-Purpose Teaching Laboratory ", IEEE International Conference on Cluster Computing, 2011, pp. 600-604.

[3] Mohammed Rashid Chowdhury, Mohammad Raihan Mahmud and Rashedur M Rahman, "Study and Performance Analysis of Various VM Placement Strategies", IEEE SNPD, June 1-3 2015.

[4] Stephen D. Burd, Gregory Gaillard, Eugene Rooney and Alessandro F. Seazzu, "Virtual Computing Laboratories Using VMware Lab Manager", 44th Hawaii International Conference on System Sciences, 2011, pp.1-9.

[5] Pooja and Asmita Pandey, "Virtual Machine Performance Measurement" , RAECS UIET Panjab University Chandigarh, 06 – 08 March, 2014.