Analysis of Network Segmentation Techniques in Cloud Data Centers

Ramaswamy Chandramouli
Computer Security Division, Information Technology Laboratory
National Institute of Standards & Technology
100 Bureau Drive, Gaithersburg, MD, USA
mouli@nist.gov

***Abstract* – Cloud Data centers are predominantly made up of Virtualized hosts. The networking infrastructure in a cloud (virtualized) data center, therefore, consists of the combination of physical IP network (data center fabric) and the virtual network residing in virtualized hosts. Network Segmentation (Isolation), Traffic flow control using firewalls and IDS/IPS form the primary network-based security techniques with the first one as the foundation for the other two. In this paper, we describe and analyze three generations of network segmentation techniques – Virtual Switches & Physical NIC-based, VLAN-based & Overlay-based. We take a detailed look at the overlay-based virtual network segmentation and its characteristics such as scalability and ease of configuration.**

Keywords-Virtual Machine; Virtual Network; Hypervisor; VLAN; Overlay-based Network; Network Segmentation

I.INTRODUCTION

Cloud data centers are computing and networking infrastructures configured for offering cloud-based services such as Infrastructure as a Service (IaaS). A great majority of servers or hosts in these centers will be found to be virtualized hosts (having the server virtualization product – the hypervisor running inside them) for reasons of scalability, agility, cost-efficiency of operations and perhaps even security. We will call these hosts as hypervisor hosts or virtualized hosts throughout this paper. Cloud data centers, because of the predominant presence of hypervisor/virtualized hosts are also called Virtualized data centers as well. A Hypervisor host has multiple virtual machines (VMs) running in each of them and in each VM one or more applications may be hosted. These applications therefore are referred to as Virtual workloads.

From the point of view of user accessibility, connectivity and security, the VMs play the same role as physical hosts in non-virtualized data centers, since as computing nodes (or endpoints), VMs house the resources that provide the functionality for one or more aspects of any application – user interface, application logic or data access. Hence it is no surprise that applications exclusively providing only one of these functional aspects are categorized as belonging to a Tier. Hence, a common architecture for applications is a 3-tier architecture consisting of Web, Application and Database tiers providing respectively the functions – user interface, application logic and data access. In other words, a VM could be hosting a Web tier, Application tier or Database tier or a combination of tiers in a data center.

From the description of the role of VMs, it should be obvious that they are the counterparts of Web Server, Application Server and Database Server in virtualized data centers, their only difference being that they run on virtual machines instead of on physical machines. Hence VMs are entities or nodes that need to be connected to each other (e.g., enable one application tier to communicate with another – A web server to communicate with Application server) and also be accessible to entities external to the data center (e.g., users accessing the application running on the VMs). The twin needs of accessibility and network connectivity, in turn, requires that each of these VMs (just like their physical counterparts) must have a distinct set of network identifiers or addresses (i.e., MAC address, IP address etc). Therefore, it goes without saying, that VMs are targets to be protected as well.

In spite of the the above common requirements between VMs and physical servers, the connectivity paradigm in which VMs are involved in brings in a different networking picture for the data center as a whole. Since multiple VMs reside in a single physical host, they may need to be connected to each other and also to the network on which the physical host itself is a node (external, enterprise or data center network). Hence there exists a capability in every server virtualization product, to define a network inside a virtualized host. This network unlike the overall data center network, is entirely software-defined and is a feature needed to provide connectivity among the VMs residing in a virtualized host as well as to provide connectivity to VMs residing in a virtualized host to the external network. This network inside a physical (virtualized) host is therefore a virtual network with its nodes (i.e., VMs) being the virtual nodes. Connectivity among the VMs (i.e., virtual nodes) and between a VM to the external network through the physical

network interface cards (Physical NICs) of the virtualized host are all enabled through another software-defined element or a set of elements called the virtual switches. The VMs themselves communicate to the virtual switches using software-defined virtual network interface cards (virtual NICs). Thus we see that a virtual network inside a virtualized host is made up of Virtual NICs and virtual switches with some communication links from virtual switches (called uplinks of the virtual switches) terminating in one or more physical NICs of the virtualized host.

The above description, together with our knowledge of the network topologies found in conventional data centers, now provides us with the picture of networking infrastructure in a cloud (or virtualized) data center - as one that is made up of the combination of the physical network (called the datacenter fabric) and the virtual network residing in each of the virtualized hosts. We all know the general network-based security techniques used for protection of resources in a conventional data center without any virtualized hosts, such as –Network segmentation or isolation, control of network traffic flows based on various parameters using devices called Firewalls and use of special-purpose network snooping devices (Intrusion Detection Systems (IDS)/Intrusion Prevention Systems (IPS)) to detect malicious traffic entering into or going out of the network. We consider these to be primary network-based security protections, although we do recognize that network services such as Dynamic Host Configuration Protocol (DHCP), Network Address Translation (NAT), Load Balancers etc do contribute to the availability aspect of the network security. Out of these three primary network-based security techniques, the network segmentation forms a foundational technique for supporting other two forms of network-based security protection. The goal of network segmentation in cloud data center environment is to enable logical separation (or isolation) among customers or tenants of (say) an IaaS cloud service.

The objectives of this paper are twofold. One is to describe to a sufficient level of detail, the network segmentation techniques available in cloud data centers whose network infrastructure are made up of physical network and virtual network components. The second objective is to analyze the security strengths and weaknesses of each of these network segmentation techniques. Based on the nature of evolution and the features they possess, we categorize these network segmentation techniques into three generations.

The organization of this paper is as follows. Since the focus of this paper is on network segmentation, we would like to provide some clarity on the concept of network segmentation in section II. In Section III, we describe and outline the strengths and limitations of the first generation network segmentation technique for cloud data centers. This solution is based on the coarse segmentation of a data center network into external, demilitarized zone (DMZ) and internal network. The segmentation of the network using the concept of virtual LANs (segmentation at the layer 2 (L2) or data link layer) and its advantages and weaknesses are the topic of Section IV. In section V, we present in somewhat great detail (compared to the other two network segmentation techniques), the network segmentation technique using the concept of network overlays and analyze in detail the advantages of this technique in terms of the security assurances it can provide. The conclusions from our analysis are presented in Section VI.

II. CLARIFICATION ON THE CONCEPT OF NETWORK SEGMENTATION

The term network segmentation in most contexts only implies logical segmentation and not physical segmentation. Physical segmentation of a data center network is enabled using physical devices such as the Top of the Rack (ToR) switches, aggregate switches, core switches and routers as well as the physical Network Interface Cards (NICs) in each of the hosts. Space availability (for housing all physical networking devices), costs (costs of equipment procurement, installation and power requirements), and finally the complexity of configuration and management limit the extent of physical segmentation that is possible in a data center. Logical network segmentation, on the other hand, requires the deployment of (is only relevant in the context of) a logical or virtual network on top of the physical network in the data center. Hence, in this paper, the network segmentation approach is always in the context of the underlying virtual networking technology. For example, when we talk about VLAN-based network segmentation, it is in the context of a *VLAN-based virtual network*, which is the type of the underlying virtual network.

III. DESCRIPTION AND ANALYSIS OF FIRST GENERATION NETWORK SEGMENTATION (BASED ON VIRTUAL SWITCHES & PHYSICAL NICs)

The first generation network-based solution for protection of VMs merely consisted of creating a single DMZ [1] (to act as a buffer between an enterprise's internal and external network) or a combination of DMZ and targeted network

segments. This solution uses the virtual switches inside the hypervisor and the physical network interface cards (physical NICs) of the hypervisor. A configuration for creating a single DMZ within a hypervisor host is given in Figure 1.

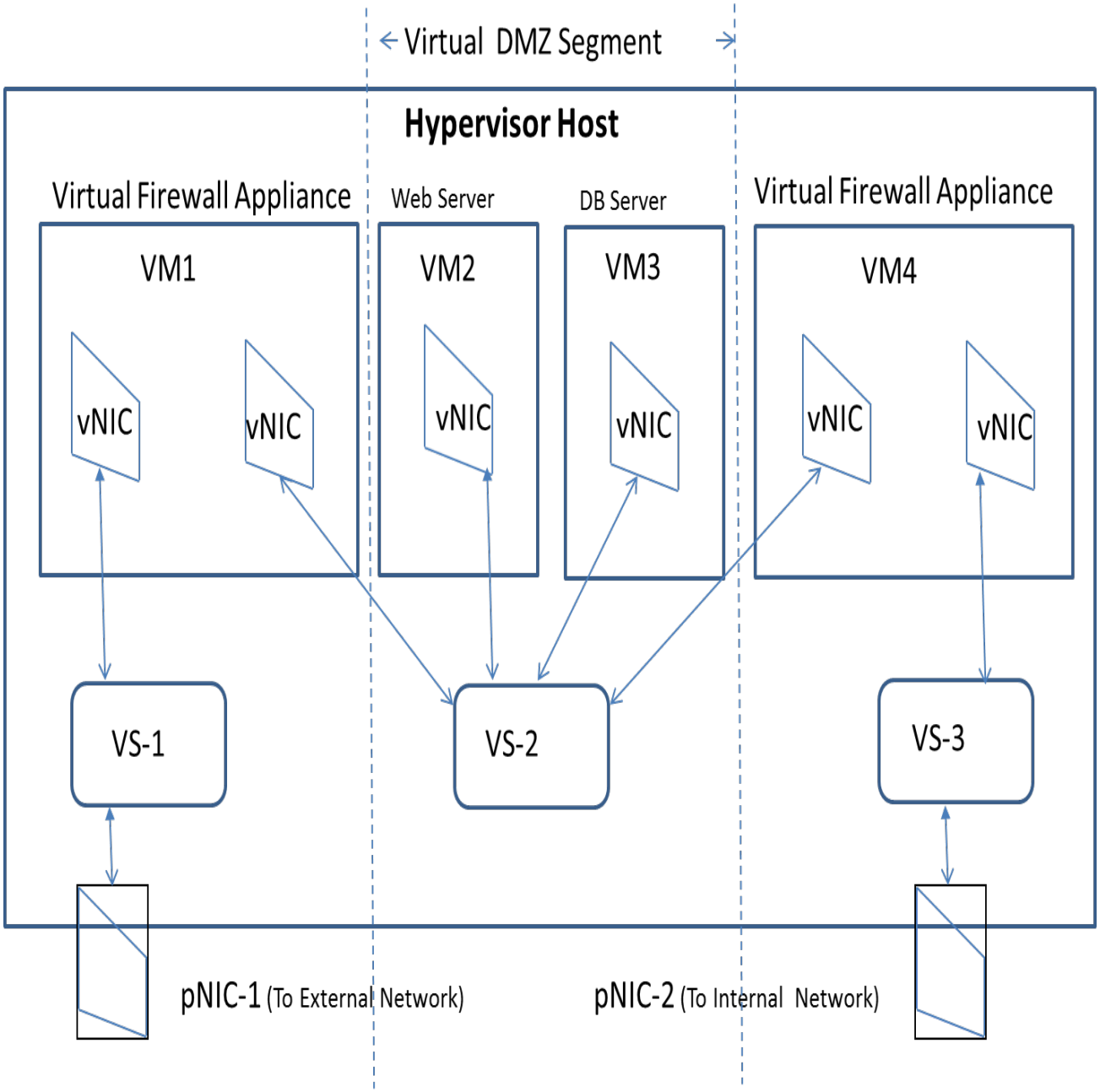

Figure 1 –Network Segmentation (Using Virtual Switches and Physical NICs)

As one can see from Figure 1, a virtual network-based DMZ has been constructed using the same architectural principles that are used to construct a DMZ in a physical network. The external firewall (between the external network and DMZ) is constructed using a virtual firewall (or a firewall appliance) running in VM1. VM1 is a multi-homed virtual server that has two virtual network interface cards (virtual NICs). One virtual NIC is connected to the external network through the physical NIC labeled as pNIC-1 via the virtual switch VS-1. The other virtual NIC on VM1 is connected to the virtual switch VS-2 to which VMs hosting a webserver (VM2) and database server (VM3) are also connected. The internal firewall in the virtual network-based DMZ (between the DMZ and internal network) is constructed using a virtual firewall hosted on VM4. Like VM1, VM4 is a multi-homed virtual server that has two virtual NICs. One virtual NIC is

connected to the internal network through the physical NIC pNIC-2 of the hypervisor host. The other virtual NIC on VM4 is connected to the virtual switch VS-2 which we have said already is connected to VMs hosting the web server and database server. Any external packet landing in VM1 through vNIC-1 can only reach the webserver in VM2 (or database server in VM3) if allowed by firewall rules of the virtual firewall hosted in VM1. Similarly any traffic from VM2 or VM3 can only reach the internal network through VM4, if it is allowed by firewall rules of the virtual firewall hosted in VM4.

In figure 1, there is only one internal logical network segment since there is only one internal-only virtual switch (not connected to any physical NIC). It is also possible to have multiple internal segments as well by using more internal-only virtual switches. The main characteristic features and limitations of this approach to achieve network segmentation are the following:

(a)Lack of Scalability: Increase in the number of logical network segments has to be achieved by increasing the number of VMs with multiple vNICs, by increasing the number of virtual switches inside a hypervisor and possibly by increasing the number of pNICs for the physical host where the hypervisor resides. Since there is limit to these values, this configuration is not scalable.

(b)Another limitation of this approach for network segmentation is configuration complexity and proneness to errors. Identical configurations have to be configured in many VMs, making the configuration error-prone.

(c)A network segment cannot span more than virtualized (hypervisor) host (since segmentation is obtained through virtual switch connectivity inside the virtualized host) – again making scalability an issue.

IV.DESCRIPTION AND ANALYSIS OF SECOND GENERATION NETWORK SEGMENTATION TECHNIQUE (VLAN-BASED)

The second generation of network segmentation for protecting virtualized infrastructure uses the concept of VLANs. A VLAN (Virtual Local Area Network) is a logical group of devices or users, grouped by function, department or application irrespective of the physical location on the LAN [2,3]. The grouping is logical and obtained by assignment of an identifier called VLAN ID to one or more ports of a switch and connecting the computing units (physical servers or VMs) to those ports. The basic objective of VLAN is logical network segmentation in order to provide broadcast containment [4]. Devices on one VLAN can only communicate directly with devices on the same VLAN and a router is needed for communication between devices on different VLANs. The VLAN implementation in virtualized hosts is enabled by using virtual switches that are VLAN aware so that VMs (as opposed to physical servers) can become VLAN end nodes. In other words, VLAN IDs are assigned to ports of a virtual switch inside a hypervisor kernel and VM assignment to those ports are made depending upon the VLAN membership of the VMs. Since in a cloud data center, VMs may belong to different consumers or cloud users, the cloud provider is thus able to provide one or more logical or virtual network segments for each tenant (for isolation of their computing/storage resources) by making VMs belonging to each of them being assigned to a different VLAN segment. These VLAN-capable virtual switches can perform tagging of all packets going out of a VM with a VLAN tag (depending upon which port it has received the packet from) and can route an incoming packet with a specific VLAN tag to the appropriate VM by sending it through a port whose VLAN ID assignment equals the VLAN tag of the packet. An example of a VLAN-based virtual network segmentation inside a hypervisor host is given in Figure 2.

The characteristics, advantages and limitations of a VLAN-based network segmentation approach are:

(a) Unlike the first generation network segmentation approach achieved using just vNICs, virtual switches and pNICs taken as a whole, VLAN configuration is based on the individual ports or groups of ports within each virtual switch, enabling a large number of virtual network segments (a virtual switch typically can support 64 ports). Further, the virtual network segments (each identified by a VLAN ID) can span more than one virtualized host.

(b) However, since the size of a VLAN ID is 12 bits, the maximum number of virtual segments possible throughout the data center is limited to approximately 4000.

(c) A VLAN implementation expects all switches (ToR and Core) to know the MAC addresses of all VMs in all VLANs. Hence in a situation where the VLAN ID to MAC address table of a ToR switch overflows, packets intended for a particular VLAN ID, may flood all links emanating from that switch to all servers instead of on just those links going into servers that hosts VMs belonging to that VLAN.

(d) A top of the rack switch (ToR) switch has many ports, at least one port for each physical server in the rack. In the simplest implementation of VLAN, every server port on the ToR switch is enabled for all VLANs. Hence the connection

linking those ports to the hypervisor host becomes a trunking link (carry traffic corresponding to multiple VLANs). Hence in the case of multicast (broadcast) messages, it becomes the responsibility of hypervisor kernel to process each of these messages for every VLAN on the network, even when the hypervisor is not hosting any active VM belonging to that VLAN [5].

(e) In a network that is designed to be aware of the presence of VMs, there will still be flooding on some ToR switch to server links. The flooding will be proportional to the number of VLANs active in that hypervisor host and the number of VMs assigned to those VLANs.

(f) There are several configuration issues that must be carefully addressed in the VLAN-based network segmentation solution. First of all, every hypervisor host is connected to an external physical switch for linking the former to the enterprise network. The VLAN configuration in this physical switch must exactly match with the VLANs configured in the virtual switches of the hypervisor host to which the physical switch is connected. Further, the links connecting the hypervisor host to these physical switches must all be configured as trunk ports (capable of supporting traffic belonging to multiple VLANs). Thirdly VMs can be migrated from one hypervisor host to another only if the source and target port group number (or VLAN identifier) is the same. In the interest of load balancing and availability, there must be a large population of hypervisor hosts available for this task and this in turn necessitates building a large VLAN (one that spans many hypervisor hosts).

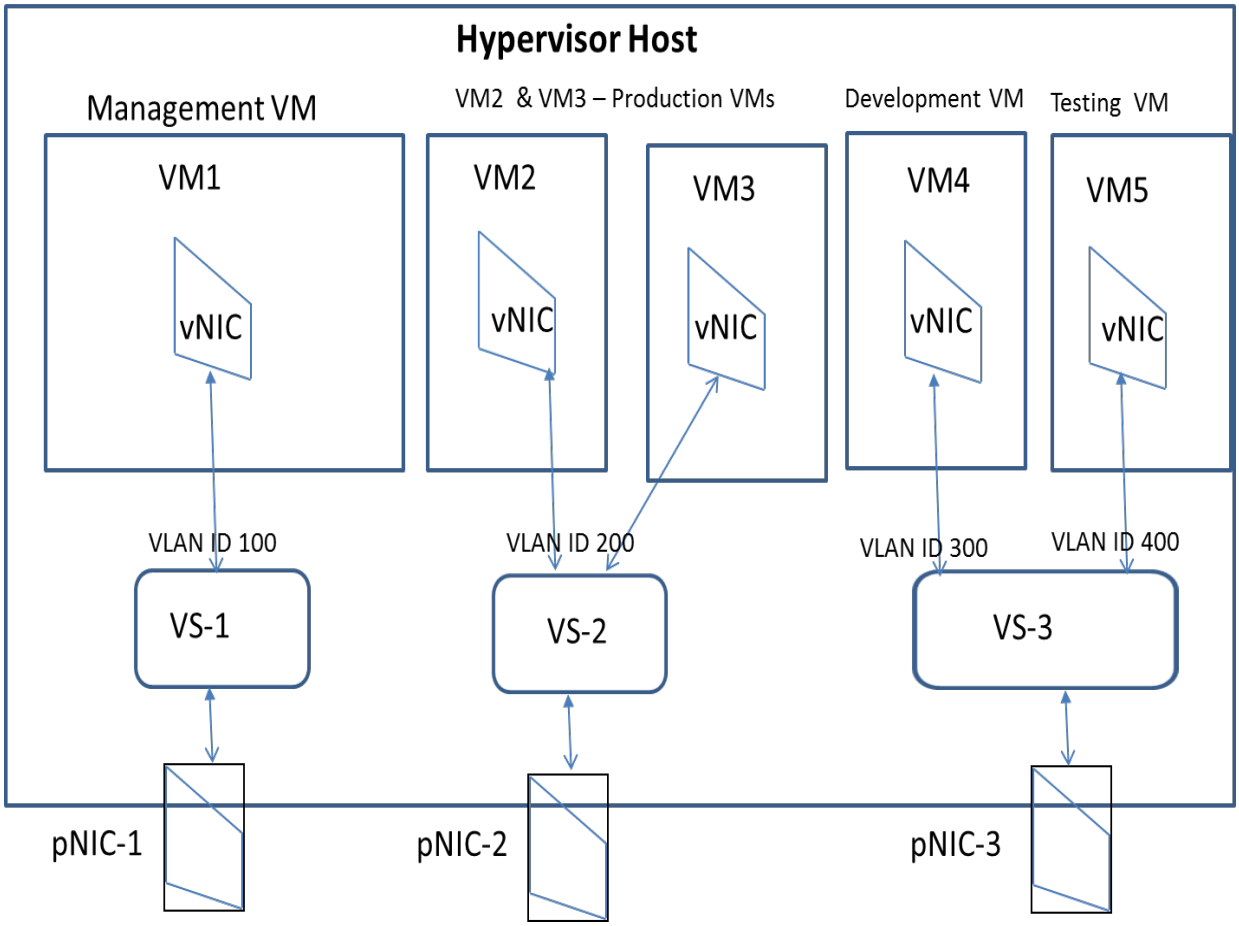

Figure 2. VLAN-based Network Segmentation

V.DESCRIPTION AND ANALYSIS OF THIRD GENERATION NETWORK SEGMENTATION TECHNIQUE

We would like to characterize the overlay-based virtual networking technology as third generation network segmentation solution. We will look at a brief description of Overlay-based virtual networking before delving into the analysis of its features and strengths

A.Description of Overlay-based Virtual Network Segmentation in Cloud Data Center

Let us briefly look at the implementation of Overlay Network and its associated segmentation in a cloud data center. The physical network topology of the data center could be either 2 or 3 layer with a reliable IP backbone. A two layer topology consists of just Top of the Rack switches and the core switches. The top of the rack (ToR) switches sit on top of a rack that contains multiple servers or hosts. Most of the hosts in our context are virtualized hosts which means they run a hypervisor inside each of them which in turn hosts multiple VMs inside each. Several ToR switches are connected to some core switches. The core switches provide connectivity to the outside world for the data center as a whole by connecting to the Internet or a VPN. In a data center with three layer network topology, there is an intermediate layer of switches between ToR switches and core switches called Aggregation switches. Aggregation switches provide connectivity between ToR switches and core switches as they are connected in a mesh topology to ToR switches (every aggregation switch is connected to all ToR switches in the data center).

In the Overlay-based virtual networking, isolation is realized by encapsulating an Ethernet frame received from a VM as follows. Out of the three encapsulation schemes (or overlay schemes) – VXLAN, GRE and STT [5], let us now look at the encapsulation process in VXLAN [6]: First, the Ethernet frame received from a VM is augmented with corresponding encapsulating protocol IDs (e.g., 24 bit VXLAN IDs) of the virtual Layer 2 (L2) segment to which both source and destination VMs are assigned. The encapsulated Ethernet frame is then further encapsulated with UDP-IP headers. The functional module that performs this encapsulation sits in the kernel of the hypervisor and is called a VXLAN tunnel endpoint (VTEP) [7]. The source IP address is the IP address of VTEP that is generating the encapsulated packet and the destination IP address is the IP address of VTEP in a remote hypervisor host sitting anywhere in the data center network that houses the destination VM. Thus, we see that VXLAN encapsulation enables creation of a virtual Layer 2 segment that can span not only different hypervisor hosts but also IP subnets within the data center. This is shown schematically in Figure 3 below:

The VXLAN based network segmentation can be configured to provide isolation among resources of multiple tenants of a cloud data center as follows. A particular tenant can be assigned two or more VXLAN segments (or IDs). The tenant can utilize the multiple VXLAN segments for assigning them to different tiers (Web, Application or Database) of the application the tenant is hosting in the data center. Selective connectivity can be established among VXLAN segments belonging to the same tenant while communication between VXLAN segments belonging to different tenants can be prohibited.

B.Analysis of Overlay-based Network Segmentation

The features and advantages of Overlay-based network segmentation are as follows:

(a) With network devices supporting programmability through standardized interfaces (being part of the SDN framework) [8], many of the network security configuration can be automated – such as the firewall rules etc

(b) Each VXLAN network identifier (VNID) is defined over a 24 bit field length rather than 12 bit field length as in the case of VLAN IDs. Hence the namespace for VXLANs is about 16 million as opposed to 4096 for VLANs. Further VXLAN is Layer 2 overlay scheme over a Layer 3 work. Hence unlike a VLAN scheme which has to use the Spanning Tree Routing Protocol to forward packets, VXLANs can use the ECMP protocol of Layer 3 [9], thus efficiently utilizing all the available network links in the network fabric of the data center.

(c) A highly scalable network security configuration is possible not only due to the unlimited availability of VXLAN IDs, but also due to the fact that the encapsulating frame is IP/UDP packet, and hence the number of virtual networks is limited only by the size of IP subnets that can be defined within the data center and not by the number of ports in virtual switches as in the case of VLAN-based network configuration. Further, by using internal, non-routable IP addresses for VMs (using DHCP and NAT capabilities) running within virtualized hosts, the number of virtual networks that can be realized is even higher.

(d) In a data center that is offered for IaaS cloud service, isolation between the tenants (cloud service subscribers) can be achieved by assigning each of them at least one VXLAN segment (denoted by a unique VXLAN ID). Since VXLAN is a logical L2 layer logical network (called overlay network) running on top of L3 layer (IP) network inside the data center, the latter is independent of the former. The consequence of this feature is that it gives the freedom to locate the computing and/or storage nodes belonging to a particular client in any physical segment of the data center network. This freedom and flexibility in turn, helps to locate those computing/storage resources based on performance (high performance VMs for data/compute intensive workloads) and load balancing considerations.

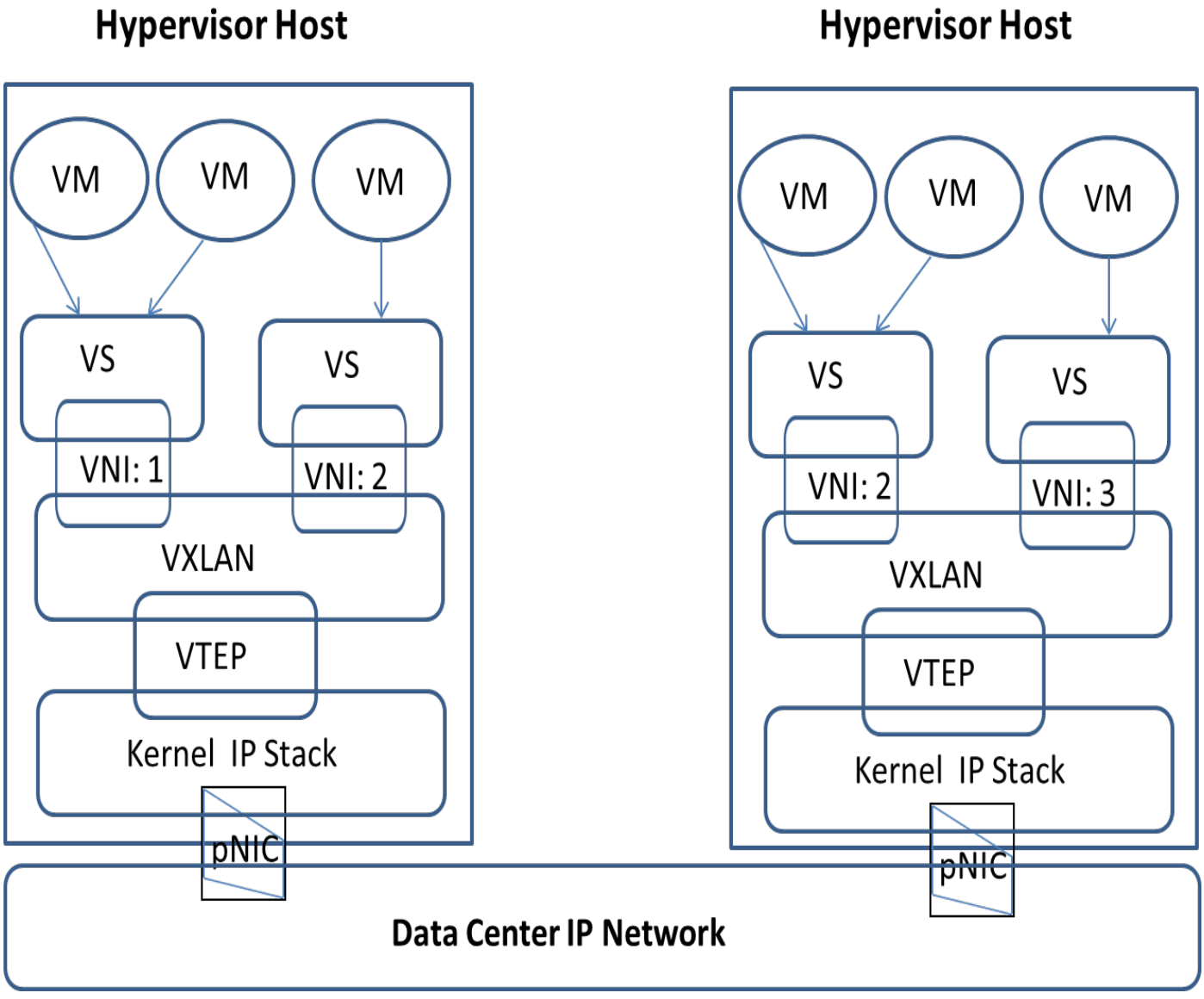

Figure 3 – Overlay-based Virtual Network Segmentation

VI.CONCLUSIONS OF OUR ANALYSIS

With the increasing adoption of cloud services by large enterprises that have to host multi-tier applications, the data center network administrators need a flexible virtual networking topology with capability to obtain the required isolation through network segmentation. At the same time, it is necessary that these virtual network segments span multiple, arbitrary IP subnets of the data center and also several hypervisor clusters. As of now, the only virtual networking technology that can provide these capabilities without a great deal of physical network reconfiguration or addition of networking resources is the overlay-based virtual networking. This degree of independence between the virtual networks and the physical networks provided by overlay-based techniques also provides the scalability and configuration ease that are needed for maintaining the logical network segmentation within large data centers. The foundational security for IaaS cloud consumer workloads thus becomes an economically and operationally viable proposition.

REFERENCES

[1] DMZ Virtualization with VMware Infrastructure, http://www.vmware.com/files/pdf/dmz_virtualization_vmware_infra_wp.pdf

[2] MAC Bridges and Virtual Bridged LANs, https://www.ietf.org/meeting/86/tutorials/86-IEEE-8021-Thaler.pdf

[3] IEEE 802.1Q Virtual LANs (VLANs), [On-line]. Available: http://www.ieee802.org/1/pages/802.1Q.html [Retrieved: June 2014]

[4] Hameed. A. Mian A.N., Finding Efficient VLAN Topology for better broadcast containment, Proceedings of Third International Conference work of the Future (NOF), Gammarth, Nov 21-23, 2012.

[5] Overlay Virtual Networking and SDDC, **http://my.ipspace.net/bin/list?id=xSDNOverlay**

[6] Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks, https://tools.ietf.org/html/rfc7348

[7]VXLAN Overview: Cisco Nexus 9000 Series Switches,] http://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-729383.pdf

[8] Open Networking Foundation, http://www.opennetworking.org

[9] Scaling Overlay Virtual Networks, http://content.ipspace.net/get/Scaling%20Overlay%20Virtual%20Networks.pdf