网络虚拟化之 NXS 概述

晓通宏志技术部：何子厚

1. VMware NSX 网络虚拟化解决方案简介

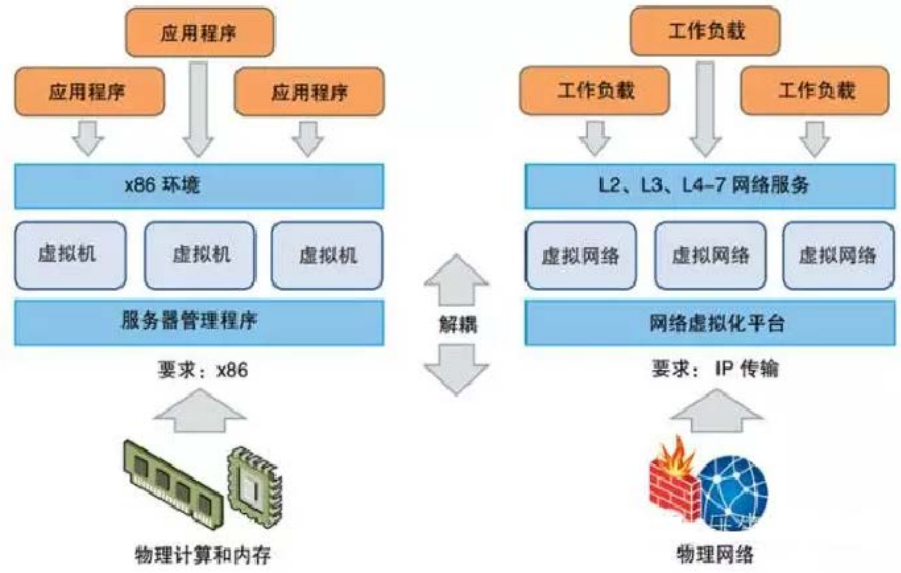
尽管 VMware NSX 网络虚拟化平台是通过收购 Nicira 而获得的，但是在收购一年多时间之后，NSX 才正式发布。在这一年多时间里，VMware 的研发人员与前 Nicira 的极客们一起通力合作，将 VMware 服务器虚拟化平台与 Nicira 网络虚拟化平台进行了融合，我们现在会发现 NSX 架构和技术细节（尤其是用于 vSphere 平台的 NSX-V）其实与早期的 Nicira NVP 平台还是有很大区别，它增加了很多 VMware 的基因在里面。

IT 行业已经直接从服务器虚拟化中获得了显著好处。服务器虚拟化解决方案降低了物理硬件的复杂性，提高了运营效率，带来了更好的安全性和冗余性，并且能够动态地重新调整底层资源的用途，以便以最佳方式快速满足日益动态化的业务应用需求。除此之外，服务器虚拟化还能节省机房空间，节省用电和制冷成本。

VMware NSX 网络虚拟化技术与 VMware 一直致力推动的服务器虚拟化技术可以进行有效的紧密结合，大大实现网络的简单化部署、配置管理。

2. 服务器虚拟化的优势移植到了网络虚拟化

以前的大二层技术一般是在物理网络底层使用 IS-IS 路由技术，再在此基础之上实现数据中心网络的二层扩展，如公有的 TRILL、SPB 技术和 Cisco 私有的 OTV、FabricPath 技术。前沿一些的网络虚拟化技术使用了 VXLAN、NVGRE 等协议，突破 VLAN 和 MAC 的限制，将数据中心的大二层网络扩展得更大。而使用 VMwareNSX，则更进一步对网络提供与对计算和存储实现的类似虚拟化功能。就像服务器虚拟化可以通过编程方式创建、删除和还原基于软件的虚拟机以及拍摄其快照一样，在 NSX 网络虚拟化平台中，也可以对基于软件的虚拟网络实现这些同样的功能。这是一种具有彻底革命性的架构，不仅数据中心能大大提高系统的敏捷性、可维护性、可扩展性，还能大大简化底层物理网络的运营模式。NSX 能够部署在任何 IP 网络上，包括所有的传统网络模型以及任何供应商提供的新一代体系结构，无需对底层网络进行重构，只需要注意将底层物理网络的 MTU 值设置为 1600 即可，因为 VXLAN 封装后的 IP 报文会增加一个头部。不难看出，VMware NSX 的核心思想其实就是将 VMware 多年致力发展的服务器虚拟化技术移植到了网络架构中，如图所示。



实现服务器虚拟化后，软件抽象层（服务器虚拟化管理程序 Hypervisor）可在软件中重现人们所熟悉的 x86 物理服务器属性，如 CPU、内存、磁盘、网卡，从而可通过编程方式以任意组合来组装这些属性，只需短短数秒，即可生成一台独一无二的虚拟机。而实现网络虚拟化后，与 Hypervisor 类似的“网络虚拟化管理程序”可在软件中重现二到七层的整套网络服务，如交换、路由、访问控制、防火墙、QoS、负载均衡。因此，与服务器虚拟化的理念相同，可以通过编程的方式以任意组合来部署这些服务，只需短短数秒，即可生成独一无二的虚拟网络（逻辑网络）。

除此之外，基于 NSX 的网络虚拟化方案还能提供更多的功能和优势。例如，就像虚拟机独立于底层 x86 平台并允许将物理服务器视为计算容量池一样，虚拟网络也独立于底层网络硬件平台并允许将物理网络视为可以按需（按使用量和用途）进行自动服务的传输容量池。对于业务或应用的激增和激退的情形，数据中心就实现了网络资源的快速分配。与传统体系结构不同，NSX 可以通过编程方式调配、更改、存储、删除和还原虚拟网络，而无需重新配置底层物理硬件或改变拓扑。这种革命性的组网方式与企业已非常熟悉的服务器虚拟化解决方案有异曲同工之妙。

由于使用了 NSX 解决方案后的逻辑网络架构产生了质的变化，以 NSX 网络平台搭建的数据中心最终达到的效果就是：无论系统规模多大，无论物理服务器、虚拟机有多少台，无论底层网络多复杂，无论多站点数据中心跨越多少地域，在 NSX 网络虚拟化解决方案的帮助下，对 IT 管理人员和用户来说，这些运行在多站点数据中心复杂网络之上的成千上万台的虚拟机，就好像是连在同一台物理交换机一样。有了 VMwareNSX，就有可以部署新一代软件定义的数据中心所需的逻辑网络。

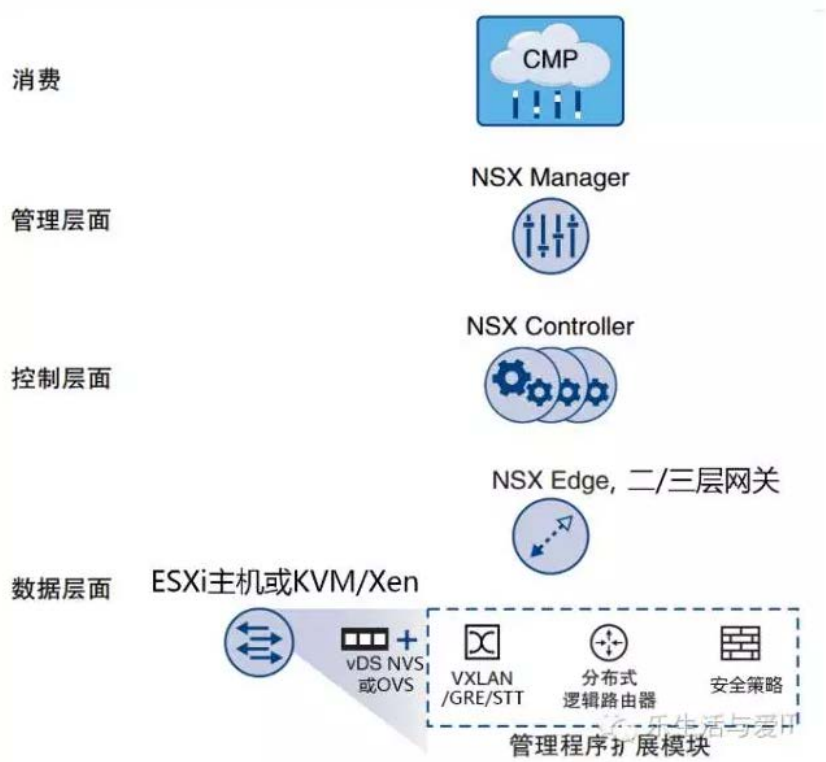
NSX 无需关心底层物理网络，那它是否一定要部署在 VMware 的虚拟化环境中？答案也是否定的。NSX 可以部署在 VMwarevSphere、KVM、Xen 等诸多虚拟化环境中，这也是 Nicira NVP 平台本来就具备的功能。

3. NSX 解决方案概览

NSX 网络虚拟化分为 vSphere 环境下的 NSX(NSX-V)和多虚拟化环境下的 NSX(NSX-MH)。这点在部署之前就需要了解，以避免错误部署。

无论使用 NSX-V 还是 NSX-MH，其基本逻辑架构都是相同的，不同点仅体现在安装软件和部署方式、配置界面，以及数据平面中的一些组件上（NSX-V 中的虚拟交换机为 vSphere 分布式交换机，而 NSX-MH 中的虚拟交换机为 OVS）。

下图是 NSX 网络虚拟化架构的基本示意图，它分为数据平面、控制平面（两个平台的分离，与提到的 SDN 架构完全吻合）、管理平面。其中数据平面中，又分分布式服务（包括逻辑交换机、逻辑路由器、逻辑防火墙）和 NSX 网关服务。控制平面的主要组件是 NSX Controller（还包含 DLRControl VM），而管理平面的主要组件是 NSX Manager（还包含 vCenter）。



NSX 数据平面主要由 NSX 虚拟交换机组成。虚拟机交换机基于 vSphere 中的分布式交换机（VDS），或基于非 VMware 虚拟化环境中的 OVS（Open vSwitch）。通过将内核模块（VIB）安装在 Hypervisor 上，实现 VXLAN、分布式路由、分布式防火墙等服务。

这里的NSX虚拟交换机可以对物理网络进行抽象化处理并直接在Hypervisor层上提供交换、路由、防火墙功能。这样有什么好处呢？首先，NSX虚拟交换机有了统一的配置界面。此外，NSX虚拟交换机利用VXLAN或STT协议实现Overlay功能，在现有物理网络之上创建一个与之解耦的独立虚拟网络，容易部署和维护。而这个虚拟网络和以前我们所熟悉的网络架构并不一样——传统的网络，不同VLAN之间地址是不能重复、冲突的，而运行在Overlay之上的虚拟网络，允许不同租户使用相同的网关或IP地址，同时保持隔离。NSX虚拟交换机连接的虚拟机是独立于虚拟网络的，就像连接到物理网络一样运行，新创建的虚拟交换机可以有效进行配置备份和还原，而它在连接虚拟机时还能实现QoS和链路聚合等诸多功能。NSX虚拟交换机还有利于实现大规模扩展，而端口镜像、NetFlow、网络运行状况检查等多种功能可以在虚拟网络内进行全面的流量管理、监控和故障排除。

数据平面还包含边界网关设备，它可以作为虚拟网络和物理网络进行通信的网关。这种设备通常是NSX Edge（NSX-V中）或二/三层网关（NSX-MH中），它以虚拟服务的形式注册到NSXManager，在虚拟网络与外界通信时，VXLAN到VLAN的转换无论发生在二层网络（桥接时）还是三层网络（路由），都可以由边界网关来进行处理。Edge、二/三层网关除了处理南北向流量外，也提供类似NFV的一些功能，如实现NAT、VPN、负载均衡等。

NSX控制平面中的主要组件是NSX Controller。它仍然是以虚拟机的形式安装，并以虚拟服务的形式与NSX Manager集成。NSXController在虚拟网络内部，可以看作是数据平面的控制单元。它与数据平面之间不会有任何数据流量的传递，只会将信令发布给数据平面，再由数据平面进行工作。因此NSXController发生任何故障都不会对数据平面的流量造成影响（这种故障不常见，因为NSXController一般都是冗余部署的）。而对外（物理网络），NSX Controller可以使用OVSDB和OpenFlow协议，作为物理网络的SDN控制器，但VMware尚未对这个功能提供官方的图形化配置界面，因此要实现这个功能，需要开发人员在API之上通过编程来实现。目前Arista Networks和Brocade两家物理硬件网络厂商的研发人员通过再开发，实现了其网络设备可以交由NSX Controller控制。

除了NSX Controller，控制平面中的其他组件还包括DLR Controller VM，用来处理三层路由协议的控制。

NSX管理平面中的主要组件是NSX Manager，可以通过NSX Manager提供的Web界面配置和管理整个NSX网络虚拟化环境的所有组件。NSX Manager提供的REST API可以为VMware高级云管理平台或第三方云管理平台（CMS/CMP）提供接口。OpenStack同样可以在这里与NSX Manager集成，使NSX与OpenStack实现融合。

NSX 可以提供以下网络服务

- **交换**：在物理网络中的任何位置实现大二层交换网络的扩展，而无需关心底层物理网络架构。
- **路由**：IP 子网之间的路由，可以完全在逻辑网络中完成。由于三层网关由 NSXController 控制并下发至所有 Hypervisor，因此流量无需经过物理路由器或三层交换机。NSX 网络虚拟化环境中的路由是在 Hypervisor 层通过分布式的方式执行的，每台 ESXi 主机的 CPU 消耗很小，可为虚拟网络架构内的路由表提供最佳路径。
- **防火墙**：安全防护可以在 Hypervisor 层以及虚拟网卡层面执行。它使用可扩展的方式实施防火墙规则，而不会像传统部署中在物理防火墙设备上形成流量的瓶颈。NSX 防火墙分布式基于 Hypervisor 内核，只产生极少的 CPU 开销，且能够线速执行。
- **逻辑负载均衡**：支持四到七层的负载均衡服务。
- **VPN 服务**：可实现二、三层 VPN 服务和 SSLVPN。
- **物理网络连接**：NSX Edge（或网关）提供虚拟网络到物理网络的二层桥接或三层路由功能。

有了 NSX 网络虚拟化解决方案，VMware 进一步完善了软件定义数据中心（Software Defined Data Center，SDDC）解决方案，即在数据中心中同时满足软件定义网络、软件定义计算、软件定义存储，并实现应用交付的自动化、新旧应用的快速创建和删除，SDDC 解决方案的核心是让客户以更小的代价来获得更灵活的、快速的业务部署、运维和管理。