数据中心交换机的“虚拟机”

—— VS （Virtual System）

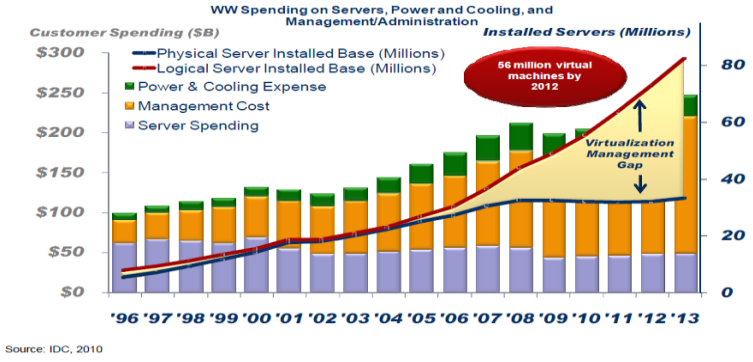
文/梁铆

导语：云计算时代，服务器虚拟化的应用如火如荼，“虚拟机”的引入提高了服务器的计算资源利用率，降低了 IT 运行维护成本；虚拟机的动态迁移，增强了服务器的可靠性、灵活性和扩展性。那么，作为云计算家族重要成员的“网络设备”，是否可以引以为镜，在网络设备上引入“虚拟机”，比如数据中心交换机的“虚拟机”，用以提升网络资源的利用率，降低数据中心运维成本，增强网络可靠性、灵活性和扩展能力?

本文旨在分析云计算时代“一虚多”的设备系统级虚拟化的必要性，并以华为新一代 VS(Virtual System)虚拟系统为例，解析其技术架构，以及典型应用场景和为客户带来的价值。

从虚拟化看数据中心交换机“虚拟机”

云时代浪潮的到来，将 IT 资源虚拟化为一片云，使得用户如同用电一样，在不必感知资源内部具体情况下，按需获得 IT 资源。“虚拟化”是云计算的关键技术，通过不同层次的虚拟化实现对物理资源的抽象，达到云资源共享并隔离的好处。根据 IDC 的分析，引入虚拟化的云以后，资源利用率由 10-15%提升到 50%-80%，更多通用的硬件被采用，资源运行和管理维护成本大幅降低，使 IT 的开销降低数十倍。



云时代的虚拟化由计算虚拟化、存储虚拟化和网络虚拟化共同组成。 就像计算虚拟化中的服务器虚拟化、桌面虚拟化等典型应用一样，网络虚拟化使得网络资源可以像计算资源一样按需供给，网络业务可以实现灵活部署并隔离，从而给云网络的客户带来实实在在的利益。

网络虚拟化在形态上分为“多虚一”和“一虚多”。 其中“多虚一”模式的虚拟化是把多个物理网络资源虚拟出一个逻辑资源，比如各种堆叠，集群技术；而“一虚多”模式相反地是把一个物理网络资源虚拟出多个逻辑资源。

“一虚多”虚拟化从应用方式上看常见的是管道虚拟化和业务虚拟化。 管道虚拟化已经在传统网络中普遍应用，网络中提供逻辑管道，依赖各种 VPN 以及 VLAN/QinQ 等技术实现对用户流量的隔离控制承载； 业务虚拟化，如 MSTP 多进程或虚拟防火墙等，通过特定业务多实例实现对业务进行逻辑隔离。 无论管道虚拟化还是业务虚拟化，都只是局部虚拟化，满足特定场景虚拟化需求的同时，网络管理者要面临多种技术的组合应用，带来部署和运维复杂性的增加等问题。为网络用户提供更简单易用的虚拟化手段，本质上需要一种彻底的系统级的虚拟化，即网络设备虚拟化，它不限于具体特定业务或管道，而是以提供整体设备级的虚拟化为目的。

从虚拟化角度，这样一种“一虚多”的网络设备虚拟化不正是和数据中心交换机的“虚拟机”不谋而合吗。

从市场驱动看数据中心交换机“虚拟机”

随着 ICT 网络特别是数据中心网络规模不断扩张，业务种类不断丰富，一方面网络管理变得越来越复杂，另一方面对于业务的隔离、安全性、可靠性等网络属性提出的要求也越来越高。此外，随着硬件能力的迅速提升，多框、集群，分布式路由交换系统的成熟，单台物理网络设备的业务处理能力已经达到一个新的高度， 如何将单台物理设备强大的业务处理能力充分利用，弹性的适应当前业务需求和未来扩展性的平滑演进呢?

具体看，客户面临的网络痛点和诉求集中在以下方面：

- ***网络设备投资成本高和设备资源利用率低的矛盾***

随着数据中心业务的蓬勃发展，不断扩大的 ICT 基础设施同时也带来了投资维护成本高昂。网络设备数量的持续增长，使得网络设备投资成本直线上升；同时，设备资源的功耗、占地空间等有关的运维成本也随着基础设施的扩充而不断增加。

与此同时，由于网络建设存在周期，为应对未来一定时期内数据中心业务的增长，客户被迫选择大于现有业务实际需求规格的网络设备，以保证网络扩展能力， 因此不可避免的出现当前网络设备使用负载不均衡，部分设备利用率低的问题。

如何化解一方面设备投资成本高昂（高 CapEx）一方面设备资源浪费的矛盾问题，是摆在客户面前的现实问题。

- ***网络设备多用户集中承载和管理隔离运维简化的矛盾***

数据中心向大规模和集中化演进的趋势，促使企业客户倾向于将不同内部和外部客户群，不同部门和组织等多个用户群的业务大集中，由统一的数据中心和网络来承载。网络设备上往往同时承载来自不同用户群的业务，如生产部门，研发部门，营销部门等，这些用户群的业务的安全、性能、可靠等网络属性往往差异很大，彼此间需要具备较好的管理隔离能力， 不同部门需要对同一设备上本部门的业务单独管理部署和维护，因此网络集中承载在带来集约化的便利的同时也面临如何对多用户群进行有效管理隔离和简化运维（高 Opex）的问题。

- ***网络设备多业务集中承载和可靠性隔离安全的矛盾***

伴随下一代数据中心的发展，网络新技术方案层出不穷，比如大二层 Trill、MAC in IP、FcoE 网络融合、各种跨数据中心互连等技术方案，同时客户希望在网络上承载的外部业务也趋向多样，这些都使得数据中心网络承载的技术和业务更加丰富，如何保证网络设备上承载的多种技术和业务彼此独立运行，减小相互间的影响，是下一代数据中心急需解决的问题。伴随着客户将更多的重要业务迁移到云计算数据中心，无疑会对网络设备可靠性和安全隔离提出更高的要求。

上述的客户痛点和诉求从市场角度驱动网络设备能够提供类似服务器“虚拟机”的能力，可以想见，数据中心交换机引入“虚拟机”后，在一台物理设备上“一虚多”出多台虚拟设备，用以提升设备资源利用率，分别用于不同用户群管理，分别承载不同业务，上述的客户矛盾将得到有效缓解。

从架构技术看数据中心交换机“虚拟机”

数据中心交换机“虚拟机”打破了物理设备资源之间的壁垒，把物理设备资源转变为可管理的逻辑资源，这些逻辑资源透明地运行在物理设备平台上，实现资源的按需分配和隔离。

为了更清晰地理解数据中心交换机“虚拟机”的价值，我们以华为 VS(Virtual System)为例，从架构技术维度理解它的内部机制和应用价值。

VS (Virtual System) 是华为 Cloud Fabric 数据中心解决方案的关键特性，提供网络设备虚拟化的技术架构，实现设备“一虚多”的虚拟化能力，即在物理设备上划分出多个逻辑或虚拟设备系统。每个虚拟系统 VS 就是设备上的“虚拟机”，可以如同一台单独设备一样独立配置、管理、维护，独立运行，承载网络业务并与其他 VS 相互隔离。 数据中心网

络通过物理设备上虚拟出来的 VS 承载不同业务或者服务于不同的用户群，达到业务隔离提升网络可靠性和安全性的目的；同时提升设备的利用率，降低用户成本；并可以实现多用户群管理隔离，有效简化运维。

虚拟化技术本质上需要实现抽象、隔离、封装。VS 在技术架构上也遵循这样的思路：

1） 抽象：

物理设备的软件系统被抽象成多个“虚拟机”，拥有逻辑独立的控制业务平面、转发平面和管理平面；硬件系统资源，包括端口、单板、内存、CPU 资源等根据用户需求被抽象成标准化的虚拟硬件，分配到不同的“虚拟机”中。经过软硬件完全抽象构建的虚拟系统具备完整的物理设备功能。

2） 隔离：

运行在同一物理设备上的多个“虚拟机”之间实现进程级隔离，抽象后的虚拟硬件以“虚拟机”为单位应用管理，不同虚拟系统之间互不影响。

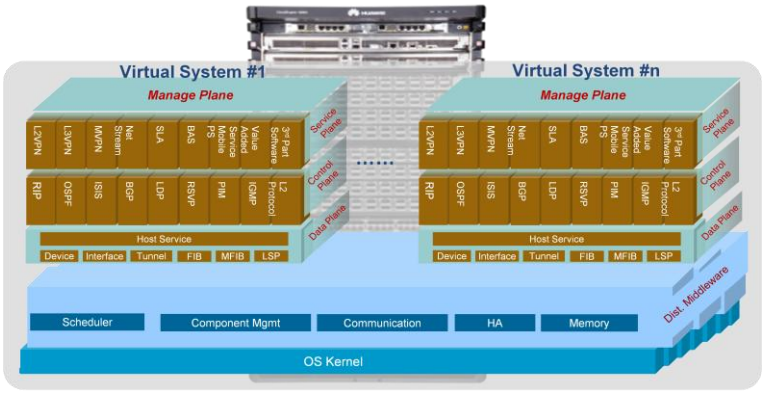
3） 封装：

将“虚拟机”封装出独立于物理位置的虚拟上下文，利用华为 VRPv8 网络操作系统的全业务分布式能力以及细粒度多进程机制，构建系统级动态迁移的能力，便于业务灵活部署，提升“虚拟机”的可靠性和设备利用效率。

VS 软件架构：

VS采用弹性细粒度分布式虚拟架构。首先，整个虚拟系统构建在网络操作系统VRPv8全业务分布式中间件的基础上，由VS控制组件实现对多个虚拟系统的统一精细的调度管理，类似服务器虚拟机中的Hypervisor；其次将设备上的网络操作系统的控制和业务平面、数据平面和管理平面进行深度的虚拟化，使得每个虚拟系统可以独立部署

业务并加载单独的配置文件，独立控制管理，例如独立复位虚拟系统，充分保证每个虚拟系统具有完全的物理设备的能力。VS进一步利用全业务分布式能力实现细粒度业务分布式部署，比如不同VS的业务模块可以分布运行在不同的单板上，有效提升硬件资源利用率。



虚拟控制和业务平面负责承载网络控制协议以及用户的业务，其可靠性和安全隔离至关重要。VS 可以允许不同虚拟系统运行在不同进程下，并精细控制进程粒度，通过进程间隔离以及独享虚拟内存空间等操作系统内核级的彻底隔离，有效地保障不同虚拟系统下的控制协议和业务的互相影响，大大增强虚拟系统的业务可靠性和安全隔离能力。同时精细的进程粒度等机制明显降低每个 VS 的系统开销，使得物理设备具备同时虚拟 16 个 VS 的扩展能力。

虚拟转发平面针对每个虚拟系统采用独立的转发环境和端口资源，各虚拟系统的数据流彻底分离，严格保障不同虚拟系统之间业务的隔离和安全性。

虚拟的管理平面为每个虚拟系统独立设置管理域，从管理用户、日志、告警到配置文件等以虚拟系统维度进行隔离，不同虚拟系统间的管理信息不可见，从根本上确保虚拟系统的独立管理能力。

VS 系统资源：

物理设备的硬件系统资源如端口、单板、内存、CPU 等被抽象成虚拟硬件划分到多个虚拟系统中，每个虚拟系统相当于独立物理设备拥有单独的硬件资源，比如当某端口划定到指定 VS 中，该 VS 将独占端口，这样一方面保证 VS 间的隔离，另一方面抽象虚拟硬件也有助于虚拟系统在设备范围迁移。

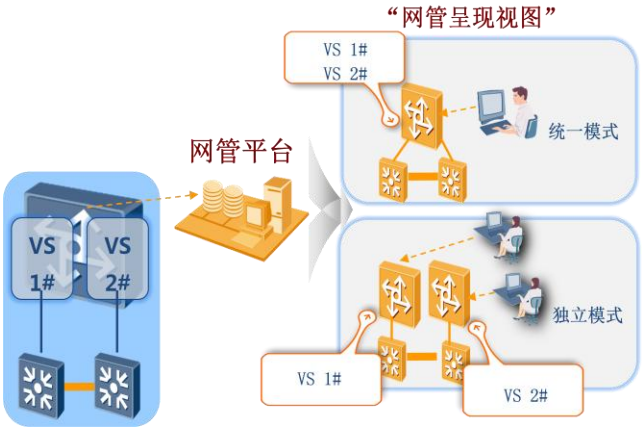
在保证虚拟系统的资源隔离的前提下，为了兼顾物理设备资源利用率，系统资源允许一定的共享，比如多个 VS 可以灵活部署允许共享物理设备的相同主控板及线卡；IPv4/IPv6 路由表、VLAN、VRF 等资源可以被多个 VS 共享，通过设定各 VS 的业务规格，保证系统资源的合理分配和使用；不同 VS 的 VLAN ID 可以重叠使用；通过逻辑接口隔离方式，不同 VS 之间可以共享同一个物理端口，节省物理链路，节约组网成本；这些都帮助物理设备上的每个虚拟系统可以弹性灵活的按需使用系统资源。

VS 管理运维：

前面已经看到，多用户群的管理运维是数据中心“虚拟机”的重要诉求，那么VS如何保证不同的虚拟系统针对不同用户群独立管理和运维呢?

VS控制管理组件以及虚拟管理平面在其中起到重要作用。每个虚拟系统在创建后，可以独立控制管理，包括按需复位、挂起、倒换、分配资源以及独立在VS视图下进行业务部署和下发配置，如同物理设备一样。每个VS可以有专门的管理员进行控制管理和业务部署，其他VS管理员无权访问本VS，方便企业不同部门独立管理本部门相关的业务。

每个VS可以进行独立运维，拥有自己独立的文件系统、配置文件、日志和告警、网管服务器等，对其他VS不可见。不同VS拥有独立网管通道和隔离的权限，充分满足多用户群共享物理设备时的独立管理以及隔离安全的需求。这种网管方式，称为独立管理模式，每个虚拟系统作为独立网元管理，有独立的网元拓扑呈现。



为灵活满足客户不同的网管需求，VS也为用户提供统一管理模式，每个虚拟系统在物理设备网元中统一管理，无独立的拓扑呈现。 统一模式适合以业务隔离为主的客户应用，网管仍统一管理，符合现有的运维习惯；独立模式则兼顾业务隔离和网管隔离的应用，网管上采用独立管理。客户可以根据需要进行自主选择。

从应用场景看数据中心交换机“虚拟机”

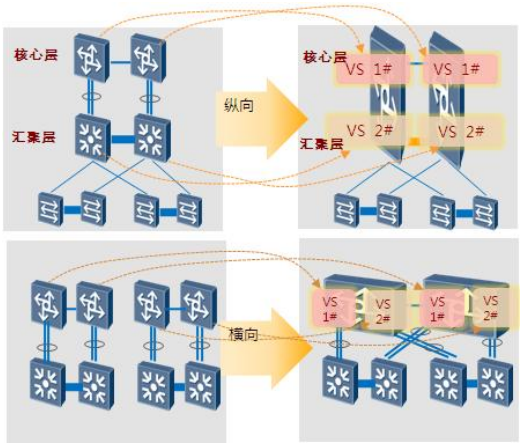
服务器通过“虚拟机”的引入衍生了很多应用，那么数据中心交换机“虚拟机”能为客户带来哪些实际的应用呢？对照前面“从市场驱动看”小节，我们从应用场景来体验VS的一些具体应用价值。

市场驱动一： 网络设备投资成本高和设备资源利用率低的矛盾

场景 1：网络节点虚拟化

按照网络节点划分虚拟系统。如通过划分为纵向的核心层和汇聚层两个虚拟系统，一台物理设备可以实现两台物理设备的组网要求；通过划分为横向并行的两个虚拟系统，虚拟化后的网络设备数量减少一半。

该应用场景带来的价值：可以在网络逻辑拓扑不变的前提下，减少物理网络设备数量，降低成本；提高设备资源利用率；降低设备功耗（电源、风扇等）、配套设备（机房、空调等）的功耗、资源的消耗；延续原有的业务体验和管理体验。

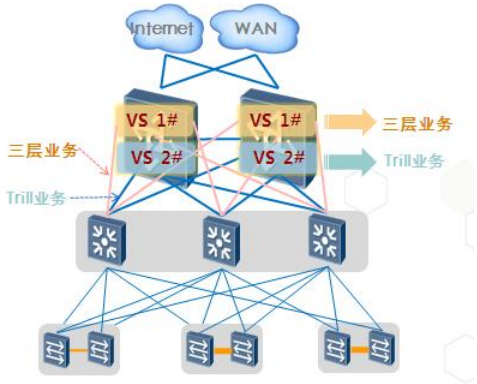


市场驱动二：网络设备多业务集中承载和可靠性隔离安全的矛盾

场景 2：业务虚拟化

按照不同业务划分虚拟系统。如 VS 1 中部署三层业务，VS 2 中部署 Trill。新业务试点存在一定不确定性和风险，通过在单独 VS 中独立部署，降低对现网其他业务影响。

该应用场景带来的价值：通过 VS 划分隔离不同业务，业务如同运行在单独设备，保障业务资源，并起到隔离安全的作用。

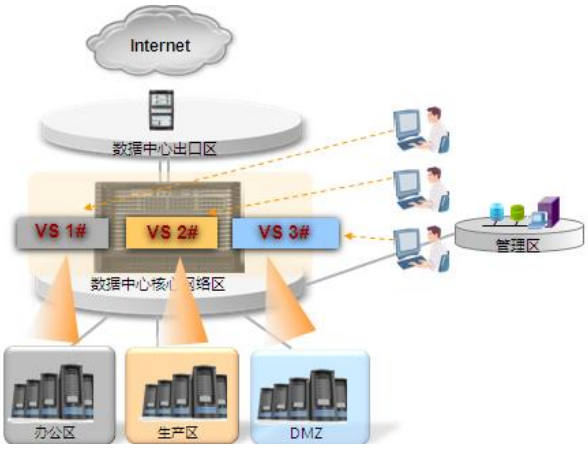


市场驱动三：网络设备多用户集中承载和管理隔离运维简化的矛盾

场景 3：用户群虚拟化

按照网络用户群划分虚拟系统。 比如按用户业务部门， 如生产、研发、市场、服务、网管等，或者按用户属性，如 Intranet、DMZ、 Extranet 等，或者按用户类别，如金融（内部办公，网银业务，信用卡业务等）进行划分。

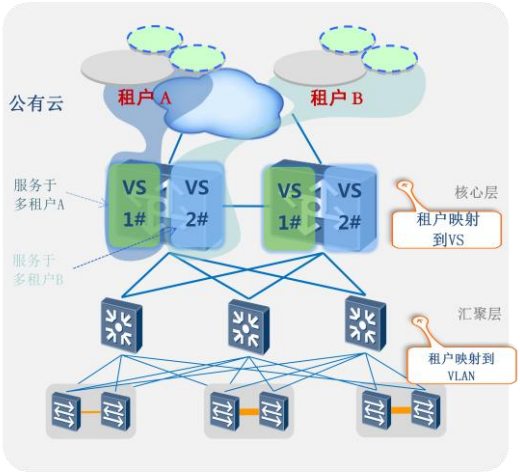
该应用场景带来的价值：不同用户群之间通过虚拟系统实现网络业务流量隔离、故障隔离，确保可靠性和安全需要；不同用户群之间通过虚拟系统实现独立网络管理，避免信息安全风险。



场景 4：大颗粒多租户

公有云中，按照不同大客户租户划分虚拟系统。比如 VS 1 服务于租户 A，VS 2 服务于租户 B，根据需要选择在核心层或汇聚层划分 VS，VS 以下层可以配合 VLAN 划分租户。

该应用场景带来的价值：相对于 VRF 隔离方式，VS 应用于多租户在灵活业务部署，在运维管理、可靠性、安全隔离等方面具有更明显优势，可以满足大型金牌级客户的高品质需求。



结束语

从虚拟化演进、市场驱动、架构技术、应用场景等几个不同维度，我们可以看到，数据中心交换机“虚拟机”对于云网络具有重要的意义和价值。华为 VS(Virtual System)作为新一代“一虚多”的设备虚拟化技术架构，将帮助客户灵活构建数据中心交换机的“虚拟机”，简化多用户管理，提升业务的可靠性及安全性，充分利用网络设备资源，降低客户成本，结合多虚一 CSS 等其他虚拟化技术，可以使网络设备随心所欲拆分或组合，提供灵活可变、可扩展的服务，把数据中心网络真正打造成伸缩自如的弹性虚拟化云网络，助力客户自由尽享云时代。