国家重点基础研究发展规划课题（项目编号：2013CB329103）
国家自然科学基金项目（项目编号：61271171）
中国科学院高新技术专项出版基金
联合资助

虚拟网络映射技术

Technologies for Virtual Network Mapping

虞红芳 孙罡 狄浩 廖丹 著

科学出版社

国家重点基础研究发展规划课题(项目编号:2013CB329103)
国家自然科学基金项目(项目编号:61271171) 联合资助
中国科学院高新技术专项出版基金

虚拟网络映射技术

虞红芳 孙 罡 狄 浩 廖 丹 著

科学出版社
北京

内 容 简 介

本书专门研究网络虚拟化环境中不同应用场景下各类虚拟网络映射问题及相关算法。本书共分7章，详细介绍网络虚拟化的基本概念和作者近年来在网络虚拟化研究中取得的一些重要成果。第1章系统介绍网络虚拟化的背景、主要概念和研究现状等内容；第2章介绍基本虚拟网络映射技术和相关算法，是本书后续虚拟网络映射的基础；第3章介绍虚拟网络的跨域映射技术，包括多域资源的统一管理机制和虚拟网络跨域映射的相关算法；第4章介绍虚拟网络的节能映射技术，给出了实现虚拟网络节能映射的算法；第5章介绍动态虚拟网络映射技术，针对资源需求量动态变化的虚拟网络请求，采用随机复用的方式以节约资源，采用动态迁移和自适应配置来应对拓扑结构动态变化的虚拟网络请求；第6章介绍单失效下虚拟网络的抗毁映射技术，包括单节点、单链路和单区域失效三种场景下虚拟网络的抗毁映射机制和算法；第7章介绍独立失效模型下可靠性保障的虚拟网络映射技术和相关算法。

本书可供从事网络虚拟化、数据中心网络和云计算研究与开发的科技人员参考。

图书在版编目(CIP)数据

虚拟网络映射技术 / 虞红芳，孙罡，狄浩，廖丹著.— 北京：科学出版社，2014.6

ISBN 978-7-03-040571-5

Ⅰ.①虚… Ⅱ.①虞… ②孙… ③狄… ④廖… Ⅲ.①虚拟网络-映射-技术 Ⅳ.①TP393

中国版本图书馆 CIP 数据核字（2014）第 092159 号

责任编辑：杨 岭 黄 嘉 / 责任校对：杨悦蕾

责任印制：余少力 / 封面设计：墨创文化

科学出版社出版

北京东黄城根北街16号

邮政编码：100717

http://www.sciencep.com

成都创新包装印刷厂印刷

科学出版社发行 各地新华书店经销

*

2014年5月第 一 版 开本：B5（720×1000）

2014年5月第一次印刷 印张：12 3/4

字数：280千字

定价：69.00元

序　言

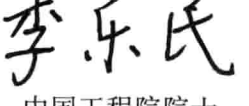
互联网作为20世纪最具影响力的技术之一，获得了巨大的成功，它改变了人们的信息交流方式。自1969年ARPANET出现以来，互联网迎来了爆炸式的发展，这40多年来涌现出的各类网络应用，对人们的工作、学习和生活产生了深远的影响。比如，电子邮件等网络应用已经极大地改变了人们的交流方式和生活习惯。

然而，互联网的日益普及和业务的快速增长和多样化，给传统的互联网结构带来了新的挑战，而传统的互联网结构在应对这些挑战时显得捉襟见肘。面对新兴互联网业务的需求，传统互联网结构暴露出了灵活性差、可靠性不够、不易管理等问题。这些问题使互联网的发展陷入了困境。

网络虚拟化技术则可以克服传统互联网所面临的一些问题。在网络虚拟化环境中，将用户提交的资源请求抽象成一个虚拟网络请求，多个虚拟网络被映射到底层物理网络上，以共享底层物理网络的资源。因此，网络虚拟化面临的一个重要挑战是如何有效地实现虚拟网络到基础设施的映射，这也成为网络虚拟化中的一个研究热点。

该书作者对虚拟网络映射问题有着长期且深入的研究。书中内容是基于作者近年来的研究成果整理而成，列出了较为全面的参考文献，免去了读者烦琐的查阅工作。该书汇聚了作者近年来在虚拟网络映射方面取得的一些重要成果，以及作者的一些研究心得，值得学习和探讨。书中用专门的章节介绍网络虚拟化的概念、应用背景、现状和发展趋势。作者提出的关于虚拟网络映射的框架机制和相关算法，对开展后续的相关研究工作有较好的借鉴作用。

总体而言，这是一本推动网络虚拟化方面研究工作的好书，特别适合从事网络虚拟化工作的研究生、教师和科技人员阅读。我愿意向有意从事网络虚拟化研究工作的读者推荐这本书。



中国工程院院士

2014年1月

前　言

由于持续增长的互联网用户群体以及不断推陈出新的网络应用，互联网体系结构面临着新的需求和挑战。虽然传统的互联网体系结构为互联网发展做出了巨大的贡献，但是其因缺乏灵活性和不便于管理等原因难以应付多样化的业务需求。因此，研究和部署全新的网络技术和结构迫在眉睫。然而，部署全新的网络技术需要更新大量网络设备的软件和硬件，这不仅会带来十分高昂的成本，而且还需要各互联网业务提供商对网络架构变化的共同认可。这些因素将会使得全新的互联网架构的出现面临重重阻力。这时，网络虚拟化技术便应运而生。

在网络虚拟化环境中，多个虚拟网络可以被映射到同一个底层物理网络上，以实现底层物理网络资源的透明共享。而网络虚拟化中面临的一个重要挑战，是如何有效地实现虚拟网络到基础设施网络的映射。高效的虚拟网络映射策略可以提高底层物理网络资源的效用，进而提高基础设施提供商的收益，同时也可以有效地降低用户的运维成本。因此，有效地实现虚拟网络的映射作为当前网络虚拟化领域的研究方向之一，有着重要的价值和意义。

作者从2009年开始从事虚拟网络映射方面的研究工作，先后承担了国家自然基金项目“基于部分合作的分布式重叠网络资源管理机制研究”和“面向多租户的虚拟数据中心可靠性设计”，以及国家973课题“服务与资源动态适配和智慧映射理论”等与网络虚拟化相关的项目。目前国内关于网络虚拟化的研究尚处在初级阶段，这方面的书籍也较为匮乏。为了给广大读者一个关于网络虚拟化的较为完整和系统的介绍，作者汇集了权威期刊和会议关于网络虚拟化的介绍，以及作者近年来在网络虚拟化方面的一些重要研究成果，著成了本书。本书系统介绍网络虚拟化的主要概念和各类虚拟网络映射的技术。希望本书能成为从事网络虚拟化、数据中心网络和云计算研究人员的有价值的读物，起到抛砖引玉的作用。

本书由虞红芳、孙罡、狄浩主编。各章编写者分别为：虞红芳(第1、6、7章)、孙罡(第4、5章)、狄浩(第2、3章)。

研究生温涛、郑少平、丁为然、苏杭、谭凌锵等参与了本书的资料整理和图表绘制工作。李乐民院士对本课题组多年的研究工作给予了很多指导和帮助，在本书的编写过程中也给予了很多宝贵意见和建议，在此一并表示衷心感谢。

最后，还要感谢国家重点基础研究发展规划课题“服务与资源动态适配和智慧映射理论(项目编号：2013CB329103)”和国家自然科学基金项目“面向多租

户的虚拟数据中心可靠性设计(项目编号：61271171)”的支持。

由于作者知识和水平有限，书中不妥之处在所难免，恳请广大读者不吝批评指正。

目　录

第 1 章　网络虚拟化概述

1.1　网络虚拟化背景

互联网在过去的 30 多年取得了巨大的成功，它的踪影已经遍布我们学习和生活的各个角落[1]。互联网能够快速发展主要得益于它优良的系统结构和网络技术，然而伴随着它推广的深入，爆炸式的增长使得自身结构的僵化问题也变得越来越明显，互联网原有的结构反而成为阻碍它进一步发展的最大障碍[2~4]。互联网是由众多服务提供商合作构建和运营支撑起来的，他们之间的利益和策略的冲突，使得对现有网络的改进只能局限于增量的更新，任何新的网络技术的引入或对现有技术的重大变革都变得异常困难。为了克服这些困难，人们做了许多努力，因此也诞生了一些新的网络技术的分支，这些都为网络虚拟化的诞生奠定了基础。

与服务器虚拟化相似，网络虚拟化旨在在一个共享的物理网络资源之上创建多个虚拟网络，同时每个虚拟网络可以独立地部署以及管理[5]。网络虚拟化概念及相关技术的引入使得网络结构的动态化和多元化成为可能，被认为是解决现有网络体系僵化问题、构建下一代互联网的理想方案[2~4]。在网络虚拟化的架构之下，用户可以根据需要定制自己的网络，摆脱硬件资源的束缚以及繁琐的网络配置过程，达到真正意义的网络作为一种服务的思想(network as a service，NaaS)。随着工业化与信息化融合进程的快速推进以及物联网、云计算等技术的规模发展，IP 网络规模、业务规模和用户规模必将进一步扩大，这也将会成为网络虚拟化发展的又一个重要的契机。除此以外，网络需求的个性化趋势使得网络虚拟化在打造智能高效的网络服务方面大有可为。

1.2　网络虚拟化研究现状

多年以来，虚拟网络(virtual network，VN)这个术语被用来描述许多基于虚拟专用网络、覆盖网络、主动可编程网络的项目。这些项目中鲜有完全的网络虚拟化方案，但都或多或少涉足其中，推动网络虚拟化的发展。表 1-1 给出了当下一些与网络虚拟化有直接或间接关系的重要项目，其中既有网络虚拟化的部署工具，也有利用网络虚拟化搭建的测试平台。它们对网络虚拟化的理解莫衷一是，实现程度也有差别。

正是因为存在着众多商业项目和学术平台的支撑，网络虚拟化的发展才会如

此迅速。现在的网络虚拟化技术已经开始摆脱各自为政的局面了，伴随着物联网、云计算等应用需求的增加，出现了一些整合了众多技术的网络虚拟化平台，其中具有代表性的当属 Flowvisor[57]、Nicira 的 NVP[53] 以及 Juniper 的 OpenContrail。

表 1-1 网络虚拟化的相关研究

项目	特征			
	涉及领域	网络技术	虚拟化层级	虚拟化程度
VNET[6,7]	虚拟机的网格计算		链路层	节点
VNRMS[8,9]	VN 管理	ATM/IP		节点、链路
X-Bone[10,11]	IP 重叠网络自动部署	IP	应用层	节点、链路
NetScript[12,13]	业务的动态组合	IP	网络层	节点
UCLP[14~16]	光路的动态配置	SONET	物理层	链路
PlanetLab[17,18]	测试平台部署和管理	IP	应用层	节点
Tempest[19,20]	备选控制架构	ATM	链路层	
VINI[21,22]	协议和业务评估		链路层	
VIOLIN[23,24]	点播增值业务部署	IP	应用层	节点
GENI[25,26]	定制 VN 的创建	混合		
Genesis[27~29]	可繁衍 VN 架构		网络层	节点、链路
Darwin[30]	资源管理和增值业务	IP		
CABO[31]	端到端增值业务部署	混合		全部
AGAVE[32~34]	QoS 感知的业务部署	IP	网络层	
Emulab[35,36]	测试平台部署和管理	IP	网络层	全部
4WARD[37]	部署管理 VN	IP	网络层	全部

1.3 网络虚拟化概况

1.3.1 发展历史

网络虚拟化概念是在最近十几年兴起的，但其技术已经有了相当的积累。在互联网发展的历史上，服务与底层设备绑定所带来的不便由来已久，人们迫切希望两者能够实现逻辑分离，在这方面也做了许多努力，其中许多都与网络虚拟化有一定的关系，按照技术规范的不同可以分为 5 大类[2~4]：虚拟局域网络(virtual local area network，VLAN)、虚拟专用网络(virtual private network，VPN)、主动可编程网络(active programmable network，APN)、覆盖网络(overlay net-

work，ON)、网络功能虚拟化(network function virtualization，NFV)。

(1)VLAN[39]是一种通过将局域网内的设备逻辑地而不是物理地划分成一个个网段来实现虚拟工作组的技术。在传统的以太网中，单一的广播域使得网络对于资源的管理手段有限。VLAN 技术的出现使得网络管理人员可以将同一物理局域网内的用户划分到不同的逻辑子网中，具有加强广播控制、简化网络管理、降低建设成本、提高网络安全等方面的作用。VLAN 技术的实现方式纷繁多样，包含基于端口、基于 MAC、基于 IP 甚至是基于用户自定义的实现方式。IEEE 802.1Q 协议统一了不同厂商的标签格式，进一步完善了 VLAN 的体系结构，也加速了 VLAN 的发展。然而，受限于标识符空间(通常为 4094 个)，传统 VLAN 已经无法满足云计算等大型网络应用场景，近年来又出现了 VXLAN[40](virtual extensible LAN)技术，可以支持多达 1600 万个 VN。

(2)VPN[41~43]是指在公用网络上建立起来的“虚拟”的专用网络(图 1-1)，它的任意两个节点间并无传统专用网络所需要的端到端的物理链路，而是构建在公用网络供应商所提供的物理网络之上，通过隧道技术实现站点间的互联，以达到共享物理网络资源的目的，所以说它是一种逻辑网络。VPN 通常用于一些组织或者公司来互连它们的子部门，也可以用于个人远端接入公司内部网络。按照互连方式可以将 VPN 划分为三大类：一层 VPN、二层 VPN 和三层 VPN，其中二层 VPN 和三层 VPN 技术已经相当成熟并被广泛应用，而一层 VPN 才刚刚起步。

图 1-1　VPN 网络模型

(3)APN 技术是将物理网络的资源通过网络可编程接口(network API)的形式暴露出来，使得用户可以自定义特定报文的处理方式。APN 的实现方式主要有两大类：第一类是利用电信技术中的信令方式将网络中的传输和控制层面区分开来，抽象出来的控制层面就可以开放网络的可编程接口，允许服务提供者控制网络的状态；第二类则是利用网络本身的资源，将控制信息封装在报文内部，路由器在收到报文时再按照其带内信息处理，达到自定义处理报文的目的。第二类在报文级粒度的处理给了 APN 更多的灵活性，更加适应复杂的网络模型。

(4)ON 是构建在已有网络上的一个逻辑网络(图 1-2)，利用隧道或封装等技术将感知节点互连起来，报文只在感知节点上处理，而在感知节点之外透明地传输。在 ON 中，虚拟的是网络的拓扑，所有的感知节点以及它们之间的联系构成了这个逻辑拓扑。ON 技术无需特定的底层网络支持，也不需要改变网络的任何特性，因此常被用于部署新的网络服务或者优化现有网络服务。

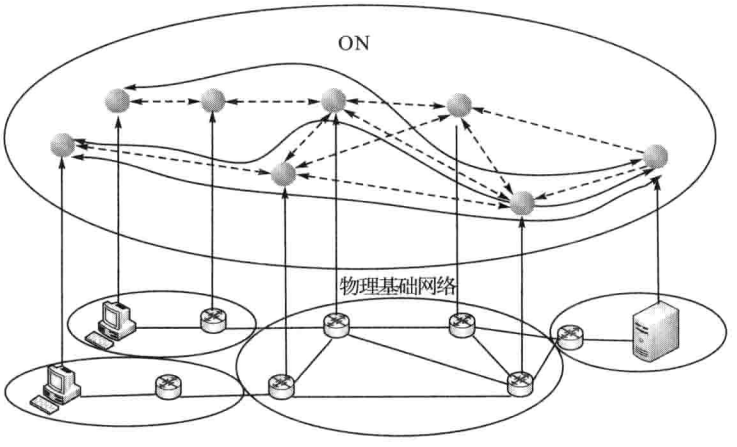


图 1-2　ON 模型

(5)NFV[44]是为简化网络部署而出现的一门技术，旨在利用已有的高性能的服务器、交换机和存储设备等代替昂贵的传统网络设备部署网络服务。原则上，该技术可运用到固定和移动网络的任何数据处理层面和控制层面。作为一个新兴的概念，其标准化的过程才刚刚开始，技术方面也有待完善，然而按照目前的发展势头来看，NFV 在 CDN(content delivery network)、IP 多媒体子系统和虚拟路由器/安全网关等领域必将大有可为。

思想和技术两者总是相辅相成、相互促进和发展的。纵观整个网络虚拟化的发展历程，从简单的想法到具体的技术，再从技术的不断发展中完善原有的思想，才有了如今庞大的网络虚拟化的技术体系。现今，对于网络虚拟化究竟该如何定义，存在着许多不同的声音[4,38]，究其缘由，多是出发点不同。虚拟化技术是为解决资源共享问题而诞生的，故笔者认为网络虚拟化的定义也应立足于此。据此，在这里给出一个当前历史下网络虚拟化的一个合理的定义：网络虚拟化是指任何用于抽象物理网络资源的技术，这些技术使物理网络资源池化，达到资源任意的分割或者合并的目的，用以构建满足上层服务需求的 VN。

1.3.2　实现方法

目前的网络虚拟化技术虽然种类繁多，但抛开外在的服务特性，其落实到底层的实现本质是基本一致的。按照网络元素的不同，可以将实现的方案分为网卡虚拟化、网络设备虚拟化和链路虚拟化三种。

(1)网卡虚拟化：网卡是物理主机的部件，同时作为网络边缘设施，负责将一台主机接入网络。一台物理主机上存在多个虚拟服务器，这会导致物理主机所具有的网卡数目少于虚拟服务器数目。因此，网卡虚拟化技术被提出用于支持一台物理主机上所有虚拟服务器的网络接入能力。网卡虚拟化技术可以分为软件实

现和硬件辅助两种。基于软件实现的网卡虚拟化通过虚拟机技术(如 VMWare 和 Xen)来实现。该技术通过软件模拟物理网卡，虚拟服务器通过虚拟网卡可以获得专用的 IP 和 MAC 地址。基于硬件辅助的网卡虚拟化技术则是通过主机内部输入/输出硬件的强化，容许建立多个连接外部设备的实例。这样，虚拟服务器就可以通过映射到该实例上实现直接接入物理网卡。

(2)网络设备虚拟化：网络设备如路由器，用于网络之中实现数据的路由(控制平面)和转发(数据平面)操作。网络设备虚拟化可以分为基于操作系统、基于控制平面和基于硬件分离这三类。基于操作系统的网络设备虚拟化也是通过虚拟机实现，在虚拟机所承载的操作系统上安装相应的网络设备软件，如路由协议。OpenFlow[45]作为基于控制/数据平面的网络设备虚拟化技术的代表，受到了广泛的关注。OpenFlow 的控制平面虚拟化是通过一个集中式的控制器来实现，而数据平面的虚拟化是通过一个共享的流表(flow table)来实现。基于硬件分离的网络设备虚拟化则是通过硬件支持，在一台物理网络设备上实现多个逻辑网络设备。

(3)链路虚拟化：通过控制数据的传输虚拟数据通道，实现链路的虚拟化。链路虚拟化可以分为隧道和标签这两类[46]。基于隧道的虚拟化是将一个数据包封装在另一个数据包中，并在网络上传输，令被封装的数据包犹如在使用专用线路。基于标签的虚拟化是通过为数据包添加标签进行定义，使网络设备可以正确地处理该分组，并指定数据包按照特定的物理路径传输。

1.3.3　商业模型

无论从思想上还是技术上，网络虚拟化都已经相当成熟，而且作为一个系统的技术集合，其架构模型也愈见清晰。在网络虚拟化的商业模型中，传统的互联网服务提供商(internet service provider，ISP)一支独大的现象被打破，取而代之的是基础设施提供商(infrastructure provider，InP)、VN 提供商(virtual network provider，VNP)和服务提供商(service provider，SP)三个实体[2,37,47]，如图 1-3 所示。

(1)InP：拥有和负责管理物理基础设施，提供用于支撑网络虚拟化的真实的底层网络资源。生产环境中，广域的物理网络中可能同时存在多个 INP，它们可以合作起来构成一个更大的 InP，当然也可以通过 VNP 协调，从而满足更大范围的网络需求。

(2)VNP：负责将 InP 提供的物理网络资源抽象成 VN 资源，并将其提供给上层的 SP。VNP 是商业模型中承上启下的中间模块，一方面，它将底层的物理资源抽象化，完成了各个分散的物理网络到一个完整统一的网络系统的转变，从而完成全局意义上的协调网络管理和优化资源利用的任务；另一方面，它可以整

合上层的网络需求，在满足用户需求的情况下，最大化网络的利用率。需要指出的是，一个较大的 VNP 可以将其管理的虚拟资源再租借给其他 VNP，充当这些 VPN 的 InP。

(3)SP：利用 VN 资源提供具体的服务，这种服务既可以是构建在网络的应用增值服务，也可以就是网络服务本身。

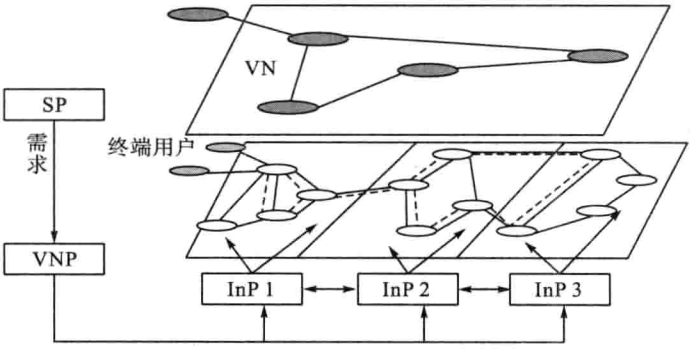


图 1-3　网络虚拟化商业模型

此外，VN 的服务部署成功之后，仍然需要对外提供一些接口，供网络的使用者进行维护。这些外部接口对应的接入者被定义为终端用户(end user)[47]，他们通常以服务访问者的身份远端接入 VN，自己本身却不属于这个 VN，所以服务提供商还需要构建额外的接入认证体系。

1.3.4　关键问题

网络虚拟化从出现至今，经过数十年的发展，虽然已经取得了很大的进展，但是由于网络虚拟化涉及很多方面的问题，因此，一方面现有的研究仍显得不够深入，有待完善，另一方面还存在一些问题亟待研究。本小节对网络虚拟化的关键研究问题归纳叙述如下[2,3]。

(1)接口(interface)。服务提供商们(SPs)综合来自一个或者多个基础设施提供商(InPs)的物理资源来构建自己的 VN，因此每个 InP 必须提供事先定义好的接口，使得 SPs 能很好地与他们之间进行需求信息的沟通。同样，终端用户(end users)和 SPs 之间以及 InPs 之间的接口也需要进行标准化定义。

(2)信令与引导(signaling and bootstrapping)。在建立 VN 之前，SPs 必须事先建立与 InPs 之间的信令链接，以传达自己的业务请求。因此，在建立 VN 之前，用于信令传递的网络连接本身是一个先决条件，只要网络虚拟化环境(network virtualization environment，NVE)还没有成熟到可以提供实现自身的信令传递的连接，信令就只能通过其他的通信方式来传递，如通过当前的互联网来传递信令。此外，还必须具备一定的引导能力，使得 SPs 能通过定制虚拟节点和虚

拟链路来建立自己的 VN。

(3)接入控制(admission control)。为了确保满足 QoS 或 SLAs(service level agreements)的要求，InPs 必须保证不会过量地为 SPs 分配资源。因此，InPs 必须得进行精确统计并运行接入控制算法以保证分配给 VN 的资源不超过底层网络的物理资源容量。此外，为了在所有分布式 VN 中避免出现违反约束条件的情况，就必须启用分布式预警机制，以确保 SPs 不能直接或者间接地出现超出分配给他们的资源量的“超负荷”情况。

(4)VN 映射(virtual network mapping)。在 NVE 中，由于 VN 的一个虚拟节点可以被映射到底层物理网络的任意一个物理节点，且一条虚拟链路可能会对应于底层物理网络的多条物理链路，所以对任意给定的 VN 而言，就存在多种向物理网络映射的方案。为了最大化共存的 VN 的数目，如何把业务提供商的 VN 请求映射到物理网络上，就显得非常重要。带有节点和边约束的 VN 映射可以被证明是一个 NP-难问题，即使是在事先给定 VN 请求的情况下也是如此[48]。目前关于 VN 映射问题的研究主要分为三大类：基本的 VN 映射问题、跨域(InP 管理域或数据中心)VN 映射问题和可靠 VN 映射问题。

(5)资源调度(resource scheduling)。在创建 VN 时，SPs 往往要求为虚拟节点和虚拟链路所分配的资源量提供特定的保障。具体而言，SPs 要求虚拟节点的 CPU 资源、磁盘和内存需求量的下界得到满足；对于虚拟链路，则希望在从尽力而为(best effort)的服务到固定丢包率和专用物理链路的延迟特性这些方面得到保证。在满足 SPs 的这些特定的资源需求的前提下，为每个 SP 建立资源独占、相互隔离的 VN，就需要有适当的资源调度策略和算法来对 InPs 的资源进行调度管理。

(6)隔离(isolation)。网络虚拟化中，多个 VN 同时运行于同一个物理网络，使用隔离保证各个 VN 之间相互不干扰。根据目的的不同可以将网路虚拟化的隔离分为安全隔离和资源隔离。安全隔离指的是各个 VN 可以运行特有的路由协议和其他服务，彼此之间不干扰，某个 VN 中由于错误配置或者网络攻击导致的故障不会影响到其他的 VN。而资源隔离是指各个 VN 获得预约的资源(CPU、内存、缓存、带宽)，这些资源不会被其他的 VN 占用，即使在其他 VN 资源耗尽时仍然能正常运行。

(7)安全(security)。网络安全是一个很大的话题，有很长的研究历史和很多研究成果，网络虚拟化的自身特点也给网络安全引入了一些新的问题。隔离在一定程度上起到安全保护的作用，能够把故障和攻击隔离在某个 VN 或者某个虚拟设备中，但是无法避免现有的对物理设备和 VN 的威胁、入侵及攻击。网络虚拟化的参与者包括 InP、SP 和 User，这三个参与实体数量众多，其中不乏为了自身利益的恶意行为，这就使得网络基础设施、VN 和用户都可能成为被攻击的对象。因此，网络虚拟化所要解决的安全问题就是分别对这三个实体的保护问题。