

基于网络功能虚拟化的网络试验平台的设计与实现

陈 鸣^{1,2)} 陶小妹²⁾ 胡 超²⁾ 许 博²⁾ 邢长友²⁾

¹⁾(南京航空航天大学计算机科学与技术学院 南京 211106)

²⁾(陆军工程大学指挥信息工程学院 南京 210007)

摘 要 新型网络设备和新协议进入实用前,最为关键的一步是在真实网络环境下进行彻底地测试验证,通常要购置大量的网络设备搭建接近真实的网络测试环境.然而,对于诸如空天地一体化网络等大型基础设施而言,进行这种测试会面临成本巨大、周期漫长、技术不成熟和不安全等诸多困难.虚拟化技术将物理资源转换为逻辑或虚拟的资源,并使得运行在抽象层之上的用户、应用或管理软件能在不需要掌握底层资源物理细节的条件下管理和使用这些资源.网络功能虚拟化(Network Functions Virtualization, NFV)基于虚拟化技术用软件可以实现网络功能,以代替传统硬件加软件方式实现各种网络功能或网络设备.这些技术为解决上述难题提供了可能.该文利用了NFV具有的功能真实、性能适中和成本低廉等特点,针对网络测试平台面临的重大技术难题提出了测试保真性的概念.测试保真性是指网络试验平台能够提供测试结果的真实性、与网络环境的兼容性和应对多种应用场景的适应性的统称.该文提出了基于网络功能虚拟化(NFV)的网络试验平台(NFV based Network Testing Platform, NFVNTTP)的体系结构,它包括了虚拟化平台功能和定制虚拟化网络功能组件两个层次.该文还提出了NFVNTTP组件模型,其中包括虚拟化路由组件、虚拟交换机组件、虚拟化网络性能组件和特种网络等几种主要组件,最后给出了一种设计NFVNTTP的系统方法.为了便于用软件配置实现或重现NFVNTTP,该文还讨论了NFV网络的形式化描述技术.为了验证本文提出的概念与方法,该文构建了一个NFVNTTP原型系统.为验证NFVNTTP原型系统的有效性,该文给出基于NFVNTTP的两个试验例子:一个试验用于对工作在空海环境中的某待测网络设备进行了兼容性测试,在该虚拟空海环境中包括了窄带数据链和同步通信卫星信道;另一个试验对两个软件定义网络(SDN)经广域网互连进行了性能测评,一个实际的SDN网络中的视频能够通过NFVNTTP传输到另一个实际的SDN中.试验结果表明,NFVNTTP不仅能够提供类似实际网络的测试保真性,而且具有经济、高效、安全等优良的试验特性.

关键词 网络功能虚拟化;试验平台;空天地网络;保真性

中图法分类号 TP393 **DOI号** 10.11897/SP.J.1016.2018.02016

Design and Implementation of Network Testing Platform Based on Network Function Virtualization

CHEN Ming^{1,2)} TAO Xiao-Mei²⁾ HU Chao²⁾ XU Bo²⁾ XING Chang-You²⁾

¹⁾(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106)

²⁾(College of Command Information System Engineering, Army Engineering University, Nanjing 210007)

Abstract Before new network equipment and protocols being entered a practical stage, the most critical process is to test and validate them in the real network environment thoroughly. Usually, a lot of network equipment should be bought to build a near real network test environment. However, for the large-scale infrastructures such as the space-air-ground integrative networks, there are many difficulties such as huge cost, long period, immature technologies and insecure and so on. Virtualization turns physical resources into logical, or virtual, resources. It enables

收稿日期:2016-06-05;在线出版日期:2017-10-20. 本课题得到国家自然科学基金(61379149)和国家“八六三”高技术研究发展计划项目(2012CB315806)资助. 陈 鸣,男,1956年生,博士,教授,博士生导师,主要研究领域为网络功能虚拟化、网络性能分析与建模、无人机网络、未来网络. E-mail: mingchennj@163.com. 陶小妹(通信作者),女,1991年生,硕士研究生,主要研究方向为网络功能虚拟化、软件定义网络. E-mail: txmyq2016@sina.com. 胡 超,男,1984年生,博士,讲师,主要研究方向为网络体系结构、分布式计算. 许 博,男,1980年生,博士,副教授,主要研究方向为网络测量与网络应用. 邢长友,男,1982年生,博士,副教授,主要研究方向为网络测量与未来网络.

users, applications, and management software operating above the abstraction layer to manage and use resources without needing to be aware of the physical details of the underlying resources. The virtualization technology based network functions virtualization (NFV) can realize various network functions or network equipment using software only instead of the traditional hardware plus software methods. These technologies provide possibility to settle the above difficult problem. To settle the outstanding technical problem for network test platform, this paper proposed the concept of the test fidelity and utilized the characteristics of NFV such as full real functions, moderate performance, low cost and so on, where the test fidelity is referred to collectively the reality of test results provided by the network testing platform, the compatibility with the network environment and the adaptability to multiple application scenarios. The architecture of network functions virtualization based network testing platform (NFVNTP) was put forward, in which there are two layers both the virtualization platform and the custom-built virtualization network function components. The components model of NFVNTP was advanced, which includes several main components such as the virtual routing component, virtual switching component, virtual network performance component, and special networks. And a method of designing NFVNTP was given. To configure and implement or rebuilt a NFVNTP, this paper also discussed a formalized description method of NFV networks. This paper built a NFVNTP prototype to validate the concepts and methods proposed by this paper. To verify the validity of NFVNTP prototype, two experimental examples based on NFVNTP were presented. One is the compatibility of the network equipment that works in the sea-sky environment was tested, in the virtual sea-sky environment there are both the narrow-bandwidth data-link and the synchronous communication satellite. Another is the performance of two software-defined networks (SDN) connected by a virtual WAN was evaluated, where the video from one actual SDN network can be transmitted via NFVNTP into the another actual SDN. Experimental results show that NFVNTP not only has the fidelity like actual networks, but also has economic, efficient, safe and other excellent experimental properties.

Keywords network functions virtualization; experimental platform; space-air-ground network; fidelity

1 引言

网络工程师为了扩展网络应用功能和提升网络性能,设计实现了许多新型网络应用和网络设备,网络科学家基于发现的网络行为规律,设计实现了新型网络机理和网络协议,而所有这些新型网络应用、网络设备、网络机理和网络协议在正式进入实用前,必须要进行透彻地性能评估。然而,对于诸如空地一体化网络等大型基础设施而言,网络性能评估面临的主要问题包括:(1)大规模网络具有的复杂、异构、非线性和时变性等特性,现有的理论分析方法和仿真工具可能无法准确反应其性能水平,有时评估分析的结果与真实水平差距很大,只有建立实际网络系统并施加真实的用户流量才能测试评估它们是否达到了设计指标;(2)建立大规模复杂试验网络

往往是一个成本巨大、周期漫长、技术不成熟的过程,这使得构建网络试验环境变得非常困难,甚至在经济上、技术上或安全上不可行。

在真实网络环境中测试被测对象(某种网络应用、协议、设备甚至网络等)具有模拟仿真和理论分析等网络性能分析方法无法比拟的优点。首先,在真实网络环境中运行网络设备和网络协议,能够保证试验测试结果的真实性。其次,被测对象通过网络接口与网络互连,使其置于网络环境中,能够测试它们与网络环境的兼容性。第三,运行在局域网或企业网中的各种网络应用能够为被测对象提供真实背景流量,从而测试出被测对象应对各种网络场景的适应性。我们将网络试验平台能够提供测试结果的真实性、与网络环境的兼容性和应对多种应用场景的适应性统称为测试保真性。然而,构建真实网络环境,尤其是构建大规模复杂网络环境所需的高昂成本以

及较长周期很难被接受,还由于此时因新型机理和关键技术往往尚在研发和验证中,系统失败或部分失败的可能是现实存在的,需要大量投资的试验往往难下决心.此外,实际试验存在着泄露频率、时序等关键通信参数的风险.因此,发明一种在实验室条件下能够提供测试逼真性的试验平台,具有重要的理论意义和实用价值.

近年来发展起来的网络功能虚拟化^[1](Network Functions Virtualization, NFV)技术基于虚拟化技术用软件来实现网络功能,以代替传统硬件加软件方式实现的各种网络功能或网络设备.多种虚拟化网络功能能够以虚拟机形式共存于通用服务器虚拟化软件中,并能够根据需要安装或移动到网络中的任意位置而不需要部署新的硬件设备.通过将网络功能软件化,NFV使网络功能摆脱了专用硬件的束缚,给通信运营商、网络用户带来了诸多好处:减少购置使用机会少、能耗高、功能单一、价格昂贵的专用硬件设备(例如某些专用网络安全设备或应用服务器)的费用;缩短了新技术引进的周期,从而加快网络的创新和新业务的部署;通过自动化控制来降低管理大量硬件设备的复杂性,同时也减少管理费用.因此,NFV具有的这些关键特征,至少在理论上使其适合于作为构建网络测试平台的支撑技术.

尽管NFV技术还在发展完善中,但是它具有的独特技术优势使其成为许多应用的不二选择.比如,构建用于测试网络新技术或新设备的网络试验环境或者用于演练网络攻防技术的网络靶场.使用传统方式构建基于大规模复杂网络的试验环境时,由于网络设备的数量庞大,从用户需求到网络规划、网络配置直至网络运行需要经历一个漫长的时期,少则数周,多则数月,而且很多技术在设计之处并不完善,其可行性及稳定性等都是不确定的,在这种情况下,直接构建真实系统存在较大的风险,就需要一种能够快速构建、低成本、高逼真、虚拟化的方式,用于在实施最终方案前对其可行性及稳定性等进行测试验证.而基于NFV技术原理构建虚拟网络的方式则符合上述要求,在通用的硬件资源池中用软件创建大规模虚拟网络只需几小时甚至几十分钟,该虚拟网络几乎具备实际网络的所有功能,能够与真实网络互联互通,构建成本低,测试逼真性高并且易于维护与管理.为了具体起见,本文主要研究基于NFV实现虚拟路由器、虚拟交换机及虚拟端系统这三种虚拟设备,由此构建基于NFV的网络试验平台的技术方法.

本文的目标是研究基于NFV的网络试验平台(NFV based Network Testing Platform, NFVNTP)的实现方法,使被测对象置于虚拟和真实相结合的网络环境中,从而得到真实可信的网络测试结果.基于NFVNTP,以特定的NFV组件能够构建起一个功能真实、性能可信、场景可定制的网络试验平台;被测对象能够与真实网络设备接口,采用虚实结合的方式与NFVNTP全真互动;最后,结合网络测量技术,能够分析评价被测对象及其网络环境整体的功能和性能.

本文的第2节讨论相关工作;第3节探索NFV网络试验平台的组织结构;第4节为满足构建NFV网络试验平台的需求,提出选择NFV软硬件平台的方法;第5节提出NFV组件模型的设计;第6节展示NFV网络试验平台的两个应用实例;第7节总结全文.

2 相关工作

从上述NFV的定义可知,基于NFV构建网络试验平台的功能完全等同于真实网络设备构建的系统:虚拟网络设备各层次之间将实际地运行着各种协议,真实地交换着各种格式的分组,甚至实际网络设备能够与NFV平台互联互通以及互换.因此,尽管基于NFV的技术与模拟和仿真这些技术都能用于网络系统的性能评估,但是前者与后者之间具有明显的不同,模拟侧重于软件,强调过程,先用数学模型刻画实际系统中发生的本质过程,通过给定不同的边界条件重现系统模型的特征来研究系统的行为.模拟过程中事件序列是由时间驱动动态完成的,模拟对系统的描述都实时地体现在对模型的计算结果中.显然,模拟的结果和所依据的模型与实际网络真实情况的相似性关系极大.限于目前数学的描述能力,模拟所采用数学模型仅能描述少量关键特征,难以全面反应系统的各个因素之间的关系,在某些情况下得到的模拟结果与实际系统差别很大.尽管目前被广为使用的网络模拟器如NS2^[2]、OPNET^[2]等的模拟结果与实际网络运行结果的等价性已经得到了认可,但在这些平台上开发的模拟程序的可移植性差,无法与实际网络结合,并且有时与网络真实情况差异较大,特别是当用户使用自行开发的新组件时更是如此.

仿真通常包括模拟过程,但更为侧重硬件和部分真实系统参与其中.NFV则需要在服务器硬件上

运行虚拟化软件,为运行于其上的操作系统映像提供一整套虚拟的兼容硬件,该虚拟硬件模拟了物理服务器所拥有的所有设备,因此,NFV 与网络真实设备等价,但在用软件实现这点上与模拟有相通之处.在一定的性能范围内,NFV 可与真实系统完全相同.表 1 列出了这几种技术的主要特点.

表 1 几种网络性能评价技术的比较

	模拟	仿真	基于 NFV
优点	用数学模型刻画系统主要特征,事件序列由时间驱动完成;用软件实现,成本较低.	包括模拟过程,但包括硬件和部分真实系统,某些部分更为真实.	与真实系统功能相同;用软件在宿主机上实现,使用较为方便.
缺点	难以全面反应系统特征;有时结果不够准确.	具有类似模拟的缺点;成本较模拟稍高.	有时性能难以保证.

目前有一类新型网络试验平台是基于虚拟化技术或 NFV 技术设计实现的.在这些平台中,所有主机、路由器、交换机、链路和控制器等都是真实的东西,只是它们均由软件而非是由软硬件结合实现的,它们与实际系统的功能完全一样,只是在某些情况下性能比实际系统要差.因此,在这类平台中,各种协议的交互过程、分组的收发等都是实际存在的.基于 LXC 容器技术的通用开放网络模拟器(Common Open Research Emulator, CORE)^①、基于 Docker 与 Open vSwitch 技术的集成多协议网络模拟器/仿真器(Integrated Multiprotocol Network Emulator/Simulator, IMUNES)^[3]和基于 Linux 内核的 Mininet^[4]的工作原理及功能相似.它们均是以物理服务器为宿主机,所有虚拟网络设备可以通过模板工具建立,虚拟网络设备如虚拟路由器等基于虚拟机生成.例如,基于标准的虚拟交换机 Open vSwitch 可以创建多个虚拟交换机.这些虚拟设备共享物理主机的硬件资源,建立的虚拟网络可通过宿主机的物理网卡与外部真实网络或位于其他物理主机上的虚拟网络相互通信. CORE 与 IMUNES 等模拟器表明了 NFV 技术在构建试验平台的可行性,但它们在应用的便利性、功能多样性等方面仍受限.

网络靶场(Cyber Range)已经成为进行网络空间安全技术研究、演练等工作必不可少的基础设施,受到美国、英国等军事强国的重视^[5].基于云计算、软件定义网络(Software Defined Networking, SDN)以及 NFV 等先进技术构建的新一代网络靶场将具有试验灵活、配置速度快、成本低、测试保真性高等特点,受到各国的推崇.

研究 NFV 的工作很多,但基于 NFV 构建网络

试验平台的讨论还比较少.文献[6]中提出了基于云计算技术和 NFV 技术的 Cloud4NFV 平台设计思想,从理论上分析验证了该平台的可行性,讨论了是控制/数据平面的功能都虚拟化(全虚拟化)还是仅控制平面功能虚拟化(半虚拟化),以及虚拟化实现的方式.文献[7]中主要讨论了 SDN^[8]与 NFV 技术如何与云计算平台相结合的方式.

3 NFVNTTP 的组织结构

NFVNTTP 的组织结构定义了试验平台各组件的组成关系,形成了支撑该平台的框架.本节主要讨论这一问题.

3.1 设计原则

设计 NFVNTTP 的组织结构应当满足如下原则:

(1)能够定制虚拟化网络功能. NFVNTTP 应当具有定制多种基本虚拟化网络功能和专用虚拟化网络功能的能力,例如通过将虚拟机配置成路由器,再利用 OSPF、RIP、BGP 等路由选择协议使其提供路由能力;控制特定端到端路径的传输能力,通过改变链路传输带宽、时延、丢包率等参数来调整网络传输性能;仿真多种特殊网络(如无线网络、卫星网络、容迟/容断网络等)的能力.

(2)能够支持虚实设备互联、互换. NFVNTTP 应当具有虚实网络设备互联甚至互换的能力,支持主流的网络协议交互运行;能够与真实企业网、局域网或网络设备接口连接,网络应用流量能够在定制的网络环境中流动.

(3)能够扩展网络规模. NFVNTTP 可由核心部分和边缘部分组成,其中核心部分为基于 NFV 的大规模虚拟网络,该网络承载在一台或多台服务器之上,边缘部分可以包括多个实际企业网或局域网或用户主机.核心部分和边缘部分的规模可以随需求而变化.

(4)能够支持多种网络功能软件. NFVNTTP 是一个研究开发平台,需要应用各种通用和专用的网络功能和信息服务功能.一方面,因特网上有大量的开源或共享的应用可供使用.另一方面,开发者可以自行开发专用的程序.这些程序共同作为平台的可用资源,以大大降低平台应用的开发成本和周期.

① U. S. Naval Research Laboratory, Common open research emulator (CORE). <https://www.nrl.navy.mil/itd/ncs/products/core>

3.2 NFVNTTP 的运行环境

图 1 给出了一个 NFVNTTP 的运行环境示例. 其中的核心部分是由 3 台承载着虚拟化软件的服务器互连而成, 该平台实际运行着各种网络协议, 它们构成了 NFVNTTP 的硬件支撑, 提供平台的计算、存储、网络等资源. 边缘部分则是由多个实际运行的企

业网或局域网和被测设备(如研发的设备、流量发生器、性能测量系统等)等部分组成. 被测对象如网络应用或网络应用协议通常运行在这些企业网/局域网中的端系统上, 被测设备通过真实的网络设备与 NFVNTTP 相连. 边缘网络可以通过 IP 路由器或交换机与核心网络部分进行交互.

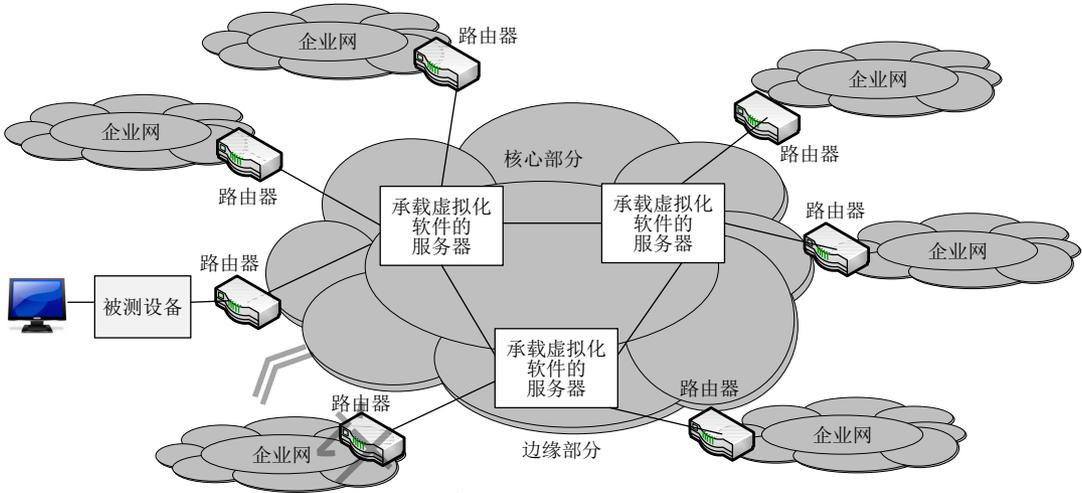


图 1 NFVNTTP 组织结构的一个示例

基于 NFVNTTP, 我们可以构造出一个大规模网络实验测试环境. 例如, 在图 2 所示的一个网络试验平台中, 核心网是在宿主服务器上构造的等级制 NFV 网络, 第一层 ISP 是由 BGP 路由器互连的主干网组成, 第二层 ISP 是由 4 个局域网组成, 它们分别是由运行 OSPF 或 RIP 路由选择协议的路由器组成, 这两部分通过虚拟路由器相连. 其中, r5、r6、r7 为 BGP 路由器, 它们互连了三个自治系统(AS). LAN1 利用 RIP 接入到 AS1, LAN2 利用 OSPF 接入到 AS2 中, 而 LAN3 和 LAN4 利用 OSPF 接入到 AS3 中. 图 2 中的 4 台移动主机 H1~H4 彼此之间可以通过 NFVNTTP 上的大规模网络进行互联互通, 其中真实网络设备与虚拟网络设备之间都真实地运行标准的网络协议, 具有实际 IP 网络的所有功

能. 例如, 一旦 r5 与 r6 之间的链路出现故障, BGP 协议将会自动通过 r7 绕行, 使得相关网络仍能保持联通. 如果 r4 与第一层 ISP 网络连接的一条链路中断, LAN4 仍能与其他局域网联通, 但若 r4 自身故障, 则 LAN4 中的用户就无法与其他局域网中的用户通信.

4 承载 NFVNTTP 的虚拟化软件

为保障虚拟化网络具有较好的传输性能, NFVNTTP 必须要承载在优选的虚拟化软件之上. 目前常见的虚拟化软件有 Xen、KVM、VMware、Docker、VirtualBox 和 LXC 等. 我们需要根据 NFVNTTP 的设计需求并结合各虚拟化软件的特点, 在量化分析评估的基础上进行优选. 例如, 假如我们初步认为 KVM^[9] 与 VMware^[10] 是比较适合承载 NFVNTTP 的两种虚拟化软件的话, 考虑到 NFVNTTP 需要承载大规模网络, 我们对 KVM 与 VMware 两者的性能进行测试比较. 在比较测试中, 我们采用的通用服务器型号为 Lenovo 30A8-A00700, 其内存为 32 GB, 磁盘 3.6 TB, CPU 为 Xeon E5-2620 处理器, 操作系统是 CentOS 6.6 64b 以及 4 端口因特网卡 F904T.

测试的主要指标是虚拟网络能够提供的传输带宽. 测试网络传输带宽的方式: 先将虚拟机配置成路

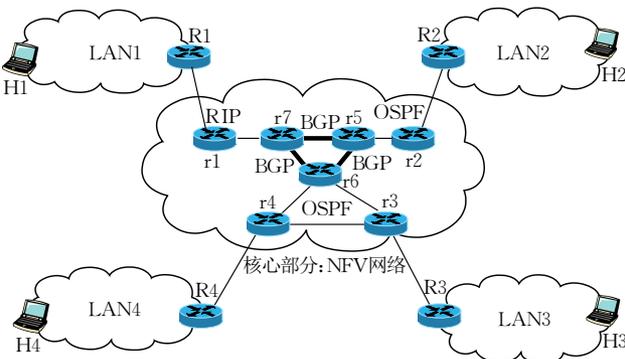


图 2 NFVNTTP 的一个实例

由器,再将多台路由器连接起来形成的端到端路径,然后测试该路径所具有的带宽.图3中显示了分别在KVM与VMware环境下,其路由器端到端路径传输带宽随链路长度增加时变化的情况.当链路为两跳时,KVM环境下带宽为10 Gbps,而VMware环境下带宽仅为1.13 Gbps;而当链路为14跳时,KVM环境下带宽为2.1 Gbps,而VMware环境下带宽仅为0.1 Gbps.显然,在相同链路长度时,KVM环境下搭建的虚拟网络具有更好的传输带宽性能.因此,在相同硬件条件下,KVM能够对NFVNTTP提供更好的仿真支持.

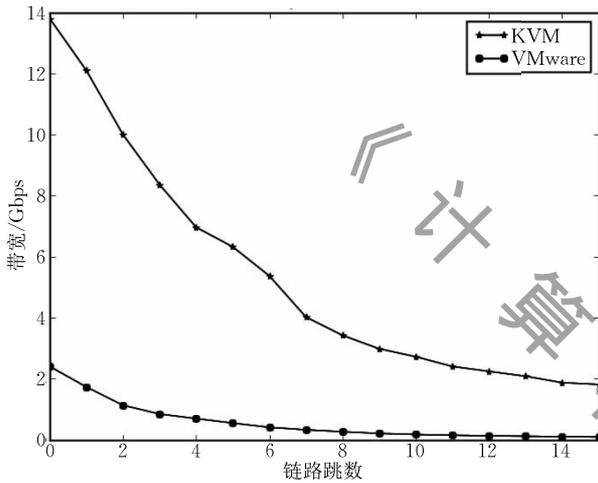


图3 KVM和VMware环境下网络带宽的比较

表2给出了两者更多特点的对比.显然,此时KVM更适合于作为支撑NFVNTTP的虚拟化软件.

表2 KVM与VMware的特征对比

	KVM	VMware
优点	是Linux内核的一部分,可随内核演进得到不断优化和改造;架构简单;性能高;开源、免费.	商业软件,技术文档齐全,市场占有率较大;技术支持便利.
缺点	对Linux内核和CPU有一定要求;资源开销较大.	架构较复杂;资源开销高,性能较低;非开源,增值开发较难;需要付费.

5 NFNTP的架构与组件模型

5.1 NFNTP架构

网络系统的体系结构是指导系统设计的高层原则,它包括网络系统各层次功能(协议、算法、机制)以及层次之间关系的集合,是指导系统设计的最高层次原则.

图4给出了网络功能虚拟化网络试验平台的架构模型.该模型包括两个部分:虚拟化平台功能和定制虚拟化网络功能组件.其中虚拟化平台功能包括虚拟化硬件平台、虚拟化软件平台和虚拟机等部分.定制虚拟化网络功能组件,需要在各虚拟机中综合利用各种IP网络协议功能模块或开发的特定功能模块,形成NFV网络和特定的网络功能,由此得到网络试验平台功能.在各种虚拟化网络组件中,路由组件和网络性能组件最为基础.

为使NFVNTTP便于使用和支持特定网络试验,该平台应当能够提供多种虚拟化网络功能组件,用户只需配置这些组件就能定制出特定的网络试验

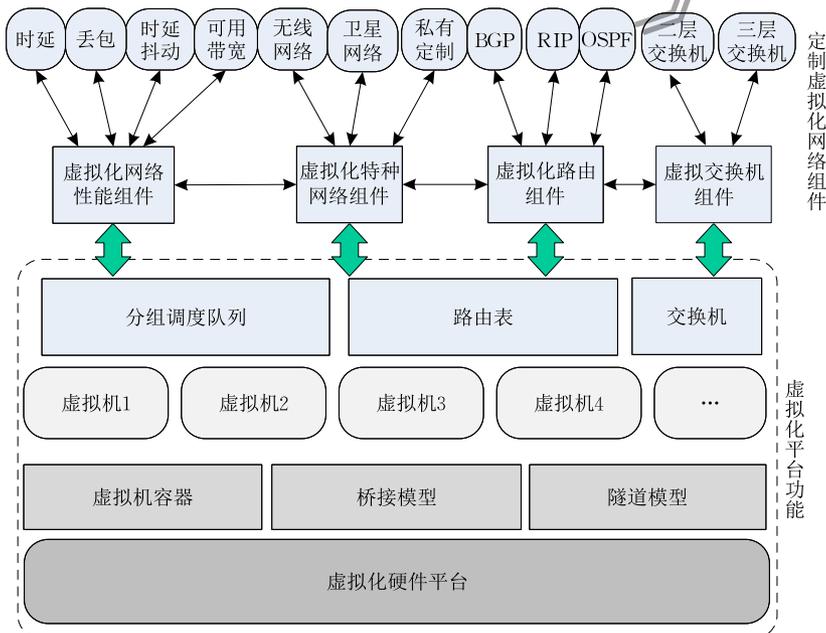


图4 NFNTP的架构

环境. 因此, 设计一系列适用的虚拟化网络功能组件是实现 NFVNTTP 的重中之重.

这些虚拟化网络功能组件应当包括设计虚拟化路由^[11]组件, 以支持 IP 组网功能; 设计虚拟化网络性能组件, 以支持网络控制性能的需求; 设计虚拟化无线网络组件, 以支持无线网络试验需求; 设计卫星网络组件, 以支持天地一体网络试验是需求; 设计容迟/容断网络 (DTN) 组件, 以支持空天网络试验, 等等.

下面简要讨论一下 NFVNTTP 的虚拟化路由组件、虚拟交换机组件、虚拟化网络性能组件和特种网络虚拟化模块等几种主要组件.

5.2 虚拟化路由组件

路由虚拟化模块的主要功能是基于路由选择协议将多台虚拟路由器配置起来以形成端到端的路径. 虚拟路由器是 NFV 的最基本单元, 设计实现该组件需要解决以下几个问题: 选择合适的路由软件部署到虚拟机中, 完成虚拟路由器配置以及路由协议配置.

选择路由软件. 路由器软件平台的种类比较多, 例如 XORP、Click、Quagga 等, 可以结合研究需要及各路由器软件平台自身的特征来选择. 支持 RIP、OSPF、BGP 等多种常用路由协议的软件路由器 Quagga 是个不错的选择.

配置虚拟路由器. 利用 Linux 操作系统对路由功能的开放性, 将 Quagga 路由软件部署到虚拟的 Linux 操作系统中, 添加 zebra 服务的配置文件并启动 zebra 服务, 此时系统进入路由器配置模式, 可像配置实体路由器一样进行各项参数的配置. 当完成对虚拟路由器各接口的 MAC 地址、IP 地址、子网掩码等相关参数的配置后, 便可将虚拟机作为路由器使用.

配置路由协议. Quagga 是一款功能较齐全的路由软件, 通过添加各路由协议的配置文件并启动对应服务, 就可对选定的路由协议进行配置, 从而可将多台路由器互连成所需的网络.

5.3 虚拟交换机组件

虚拟交换机组件的主要功能是通过学习与其相连设备的 MAC 地址来完成数据帧的转发. 该组件需要解决虚拟交换机软件的部署, 网络设备接口与虚拟交换机之间的连接以及作为 SDN 交换机时与控制器的连接等问题.

Open vSwitch 是目前被广泛使用的虚拟交换标准, 既可作为普通的二层交换机, 也可用作 SDN 交换机. Open vSwitch 的部署方式与 Quagga 的部

署方式不同, Open vSwitch 是部署到物理主机中, 而不是部署到虚拟机中, 部署成功后, 就可以创建出多台虚拟交换机 (简称 OVS 交换机), 而一台虚拟机部署了 Quagga 后只能作为一台虚拟路由器使用.

OVS 交换机作为二层交换机使用时, 在虚拟机管理工具 virt-manager 中将物理网卡或虚拟机的网卡设置为映射到虚拟交换机上, 即可完成设备与虚拟交换机的连接. 例如, 在图 5 中 ovs-br0 与 ovs-br1 为 Open vSwitch 创建的两台 OVS 交换机, 这两台交换机是互连的. eth0 为宿主机的物理网卡, 通过指令可将其与虚拟交换机 ovs-br0 绑定. 这样, 其他物理设备就可通过该网卡与 ovs-br0 连接. 类似地, 虚拟机 VM1 与 VM2 通过 virt-manager 可将虚拟网卡映射到虚拟交换机 ovs-br1 上.

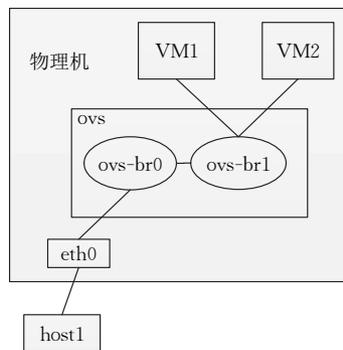


图 5 OVS 交换机与外部主机与虚拟机的连接

若要将 OVS 交换机作为 SDN 交换机使用, 先要将其与 SDN 控制器相连, 并在控制器端编写控制程序来控制交换机的某些行为. 例如, 在宿主机中执行命令“ovs-vsctl set-controller br0 tcp: <controller IP> : <port>”即可将 OVS 交换机 br0 连接到一个指定的控制器. 启动控制程序后, 可执行命令“ovs-vsctl show”查看是否连接成功.

5.4 虚拟化网络性能组件

网络性能虚拟化模块基于 Linux 中的成熟程序, 实现了对网络接口的传输性能控制, 可以定量地调节链路传输的性能.

Linux 内核 2.6 及以上版本中均包括了网络模拟功能模块 NetEm, 该模块对 Linux 的流量控制工具 tc (traffic control) 进行了功能增强, 添加了对链路延时、分组丢失、分组冗余等情况的模拟, 可用来在性能良好的局域网中模拟出各种复杂的广域网传输性能, 如产生低带宽、长时延、高丢包率等特定情况. 在本平台的虚拟机中, 可在虚拟路由器的特定接口上运行 NetEm, 从而为流经该接口的流量设置预定的路径带宽、时延、丢包率等性能参数值.

如图 6 所示,可将虚拟路由器 r2 与 r3 所有端口都设置为 100 Mbps 的传输带宽,10 ms 的时延,3% 的丢包率.设置成功后,通过在 h2 和 h3 之间使用 Iperf 等网络性能测量工具进行测量,可以检验控制的结果.经实际测试,可以发现带宽、时延、丢包率等性能参数值与设置的参数值均十分接近,从而能够达到控制网络传输性能的目的.如果该链路丢包率设置为 100%,则等同于该链路中断.

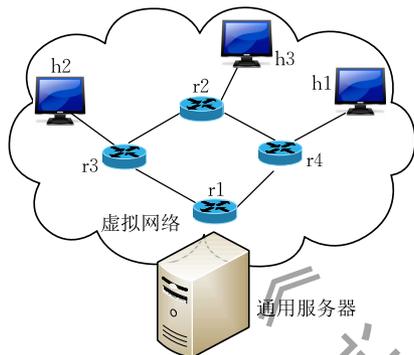


图 6 测量虚拟化网络性能组件的实例

5.5 特种网络虚拟化模块

特种网络虚拟化模块的主要功能是通过编制专用网络协议或通过控制虚拟路由器接口性能参数来模拟卫星网络、无线网络、DTN、传感网甚至水下网等特殊网络,达到对分组传输影响的目的.研发特种网络虚拟化模块需解决用网络功能虚拟化相关技术真实逼近这些特种网络的功能、性能的问题.

例如,如果要专门定制 DTN 模块,就要根据 DTN 所具有的可以在长时延、频繁中断、结点资源有限的受限网络环境下提供可靠的数据传输的特性,遵从 RFC 文档^[12-13]规定,自行开发软件或集成开源代码以实现相应的网络功能.再如,如果要定制地球同步轨道卫星通信模块,可以根据卫星通信传输的关键性能参数值,可以将其通信带宽设为 10 Mbps、往返通信时延设为 500ms 和丢包率设为 0.01%,从而达到模拟分组通过同步卫星信道的情况.同理,可以根据无线通信网络的关键性能指标,通过控制虚拟网络特定链路的传输性能以达到模拟无线网络的目的,例如,参照 IEEE 802.11 a/b/g/n 标准,可以将速度分别设为 54 Mbps、108 Mbps、150 Mbps 和 300 Mbps 并设置相应的时延和丢包率,从而模拟出所需的无线网络场景的传输性能.

6 NFV 网络的形式化描述

尽管配置与管理虚拟网络中的虚拟路由器等设

备较之配置与管理实际网络的设备简便了许多,但是对于一个大规模虚拟网络而言,无论是配置一个新网络或者重现一个已配置的网络,手工操作过程仍是一个繁琐、费时、易错的过程.因此,首先需要研究一种形式化描述 NFV 网络的方法.有了这种形式化方法,我们就能唯一、准确地描述 NFV 网络中的各种 NFV 网络视图的配置,进而可以根据该配置文件采用软件程序来实现 NFV 网络的配置.本节重点描述一种基于 XML 描述 NFV 网络的方法.

XML 的标签及层次结构可由用户根据实际情况自行定义.XML 的这种灵活性使其比较适合用于存储配置信息.本文采用 XML 文档来存储虚拟网络的路径传输性能参数信息以及虚拟路由器的配置信息.

根据虚拟路由器的端口配置及其传输性能定义等过程,将 XML 文档划分为多层,每一层对应指定的信息.文档第 1 层为根元素,指出了这是网络配置的 XML 文档.第 2 层定义了虚拟设备类型,目前主要的虚拟设备为虚拟路由器、虚拟主机等.路由器元素等还包含多个子元素,即为第 3 层的内容,其子元素标签的定义如表 3 所示,包含定义路由器名称,路由器类型,路由器通信 IP 地址/掩码,网卡信息,路由表信息等 5 项内容.第 1 层与第 2 层元素标签定义如下:

第 1 层: < NetworkConfig > ... < /Network-Config > <!-- 表示网络配置的 XML -->

第 2 层: < router > ... < /router > <!-- 路由器元素 -->

表 3 路由器标签子元素定义

子标签	含义
< rname > ... < /rname >	定义路由器名称
< category > ... < /category >	定义路由器类型
< MgtIP > ... < /MgtIP >	定义路由器通信 IP 地址及掩码
< Interfaces > ... < /Interfaces >	用于定义多个网卡的信息
< routing > ... < /routing >	用于定义路由表信息

表中用于定义多个网卡的信息这一标签中包含多个并列的网卡标签(< interface > ... < /interface >),即为文档的第 4 层,每一对网卡标签对应于虚拟路由器的一块虚拟网卡.每一对网卡标签包含多个子元素,为文档的第 5 层,这些子元素标签的定义如表 4 所示,包含接口 IP 地址/掩码,接口名称,接口类型,标记符,链路往返时延,链路丢包率,链路带宽等 7 项内容.

表 4 网卡标签子元素定义

子标签	含义
<ip>...</ip>	定义接口 IP 地址及掩码
<ename>...</ename>	定义接口名称
<etype>...</etype>	定义接口类型
<change>...</change>	定义标记符
<delay>...</delay>	定义链路往返时延
<pktloss>...</pktloss>	定义链路丢包率
<bandwidth>...</bandwidth>	定义链路带宽

表 3 中用于定义路由表信息(<routing>...</routing>)这一标签包含多个子元素,子元素标签的定义如表 5 所示,包含选路协议类型,选路协议父标签两项内容,其中选路协议父标签“<paras>...</paras>”会并列多对,每一对包含一项子元素

“<info>...</info>”,用于描述选路协议必须的参数信息,以 OSPF 协议为例:“<info>network 192.168.1.0/24 area 0.0.0.0</info>”。

表 5 路由表标签子元素定义

子标签	含义
<rtype>...</rtype>	定义选路协议类型
<paras>...</paras>	选路协议父标签

根据上述 XML 文档标签及层次结构的设计结果,可以绘出与整个 XML 文档相对应的 XML 模型图,并为每一对标签文本赋一示例内容.图 7 给出了虚拟网络配置的 XML 模型。

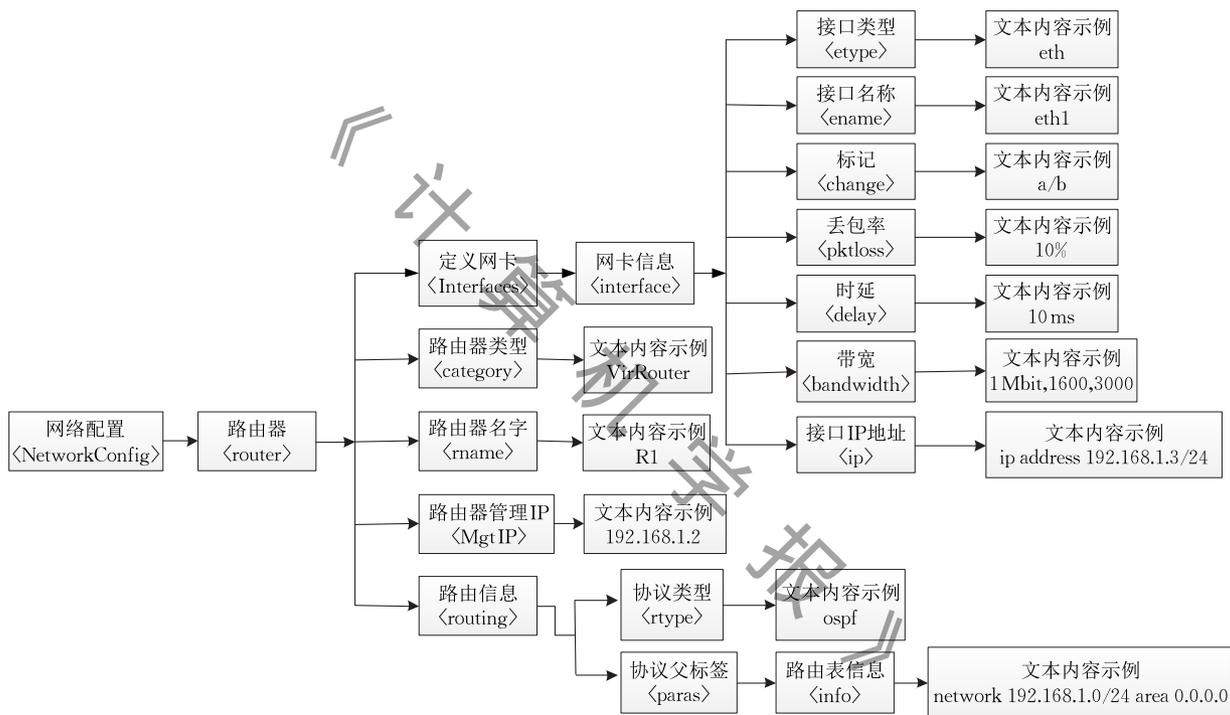


图 7 虚拟网络配置的 XML 模型

7 试验及其分析

本节先给出一个在 NFVNTTP 上设置网络链路带宽、往返时延(RTT)性能参数的简单示例及其测试结果,然后给出两个应用实例以说明 NFVNTTP 的用途和使用方法。

7.1 试验 1:网络链路性能控制

我们首先在 NFVNTTP 中利用 4 台虚拟路由器 R1~R4 连接两台虚拟端系统 H1 与 H2 的虚拟网络(参见图 8)。然后,在 H1 和 H2 上运行 Iperf 测量程序;在 R1 上运行 NetEm 模块,对其一个接口的带宽、单向时延和丢包率等传输性能进行设置;在任

何一台虚拟机上都可以安装 Wireshark,俘获实际发送的报文进行分析协议和数据内容。

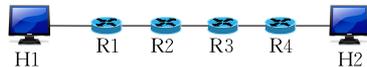


图 8 测试一个简单网络的传输性能

(1)控制接口单向时延.在虚拟路由器 R1 中启动 NetEm 功能,控制经指定接口发送的数据分组的单向时延,从而达到控制该链路单向时延的目的.例如,在 R1 中执行命令“tc qdisc add dev eth1 root netem delay 10ms”(其中“eth1”指代第一块网卡),将 R1 中名为“eth1”接口发送的数据分组的链路 RTT 设置为 10 ms.注意到 NetEm 的设置是单向

的,上述命令仅对从虚拟路由器的 eth1 口发送的数据分组有效.表 6 分别给出了单向时延、带宽和丢包率的设置值与实测值.

表 6 设置时延测试结果

设置单向时延值/ms	实测时延值/ms	设置带宽值/bps	实测带宽值/bps	设置丢包率/%	实测丢包率/%
10	10.5	64 K	61 K	10	9
20	20.6	128 K	122 K	20	20
30	30.5	512 K	489 K	30	27
40	40.6	16 M	15.2 M	40	37
50	50.7	32 M	30.6 M	50	51
60	60.6	64 M	61 M	60	61
70	70.4	128 M	125 M	100	100

(2)控制接口带宽. R1 运行 NetEm,能够控制指定接口发送数据分组的传输带宽,从而达到控制整条路径传输带宽的目的.例如,在 R1 中执行命令“tc qdisc add dev eth1 root tbf rate 1Mbit buffer 1600 limit 3000”,将 R1 中名为“eth1”接口发送的数据分组的传输带宽设置为 1Mbps. NetEm 对带宽的控制是单向的,因此,在设置时需要根据试验需求,确认传输方向或在链路或路径两端的结点接口做必要设置.表 6 给出了带宽设置值与实测值.

(3)控制接口丢包率. R1 运行 NetEm,控制经指定接口发送的数据分组的丢包率,从而达到控制链路丢包率的目的.例如,在 R1 中执行命令“tc qdisc add dev eth1 root netem loss 10%”,将 R1 中名为“eth1”接口发送的数据分组的丢包率设置为 10%,限制了 R1 中名为“eth1”的接口发送数据分组的丢包率.表 6 给出了丢包率设置值与实测值.

从以上 3 组测试结果看,用 NetEm 控制 IP 网络的传输性能,对时延、带宽、丢包率等网络性能的设置值与实际测试值十分接近,控制效果好.因此,只要无线网络、卫星网络、DTN 等特种网络的性能特征能够用典型参数值或某个函数来描述,在 NFVNTP 就可以方便地模拟出来.另外,NetEm 还能够模拟分组的时延抖动、重复、失序等多种情况,由此可见,使用 NFVNTP 作为试验测试平台在理论上及技术上都是可行和可用的.

试验 1 结果表明:在 NFVNTP 中可以方便地用 NetEm 控制虚拟设备接口的 RTT、带宽和丢包率等性能指标.试验测试结果表明,RTT、带宽和丢包率的控制效果好,实测值与设置值非常接近

7.2 试验 2:在海空环境中测评网络设备

应用示例 2 描述了这样一种场景:当某飞机在

海洋上空巡航时,它所携带的数据链设备利用低速无线信道将控制信息传输给卫星,再通过卫星将信息传输给卫星地球站,最后通过陆基有线网络传输到控制中心.某新研发的数据链设备完成初样后,需要对该新设备进行反复测试,待其功能正确、工作稳定可靠后才能正式投入生产和使用.在该数据链设备研制过程中,如果在真实环境中反复测试,不仅试验开销巨大,而且容易泄露其通信参数.因此,基于 NFVNTP 为测试该数据链设备而构建一个保真的网络试验平台具有重要意义.

图 9 示意了海空环境下的试验网络,包括实体网络和虚拟网络两大部分.图中左下部分包括了被测数据链设备的实体网络,主机 H2 用于发送各种测试报文,通过通用服务器的 eth1 网卡与 NFV 虚拟网络连通.图中右下部分是包括测试目的局域网和目的主机 H1,测试报文通过虚拟路由器 r1 映射到服务器物理接口 eth0,再与 LAN1 中的实体路由器 R1 相连,最终呈现在 H1 屏幕上.其中:(1)端系统主机 H1 和 H2 采用 Lenovo T440p 便携机,分置于端到端路径的两侧;(2)被测设备为数据链设备,置于被测网络回路中;(3)在 NFVNTP 上按实际陆基 IP 网络拓扑设定多个 BGP 路由器,将该网络划分为多个自治系统(AS).虚拟路由器 r4、r5、r6 配置为 BGP 路由器,整个网络划分为三个 AS.

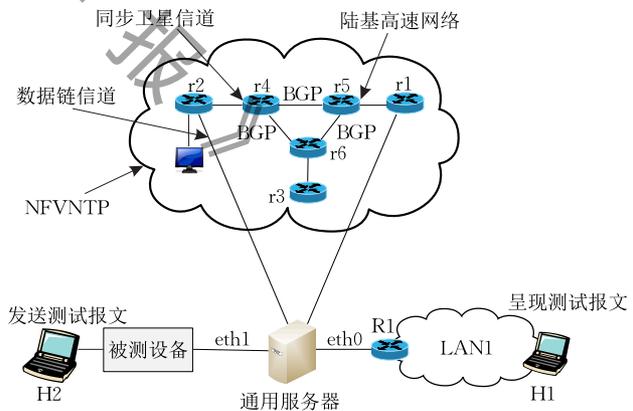


图 9 在海空环境下的试验网络中测试数据链设备

AS1 用来模拟无线网络,r2 与 eth1 的连接模拟数据链无线信道,使用 NetEm 控制通信信道的速率为 200 kbps,丢包率为 5%,时延为 30 ms;r4 与 r2 的连接模拟同步卫星信道,使用 NetEm 控制通信信道为 10 Mbps,丢包率为 1%,时延为 500 ms.其他 AS 模拟了高带宽的陆基网络,速率为 1 Gbps,时延为 30 ms,丢包率为 0.000 01%.

试验 2 结果表明:(1)NFVNTP 可以软件定义产生大规模网络,该网络的功能等同于实际网络;(2)在 NFVNTP 中可以同时模拟多种特定网络,并考察这些网络综合作用的后果;(3)利用 NFVNTP 可以将被测设备实际置于网络回路中,测试该设备的物理接口和逻辑功能.因此,NFVNTP 能够为测试数据链设备或通信装备提供较理想的保真测试环境.

7.3 试验 3:OpenFlow 网络通过广域网互连

OpenFlow^[14]网络是符合软件定义网络(SDN)南向接口标准的网络,它分离控制平面和数据平面的工作方式网络创新提供了新型体系结构.目前实践中广泛采用利用互联网将网络边缘的多个基于 OpenFlow 的数据中心网络互连起来^[15].本试验研究了利用 NFVNTP 互连两个 OpenFlow 子网的过程和方法.

首先采用 2 台型号为 Pica8 P3297 的 OpenFlow 交换机、2 台 Linux 的 Lenovo T4900v PC 运行 POX 控制器和 4 台 Windows 主机 H11、H12、H21 和 H22 按图 10 所示方式构成两个 OpenFlow 子网. NFVNTP 仍采用与图 9 中相同的广域网.这两台 OpenFlow 交换机的一个端口分别与通用服务器的一个物理接口相连,即 OFS1 端口 1 通过服务器 eth1 接口与虚拟路由器 r1 相连;OFS2 端口 1 通过服务器 eth0 接口与虚拟路由器 r2 相连.

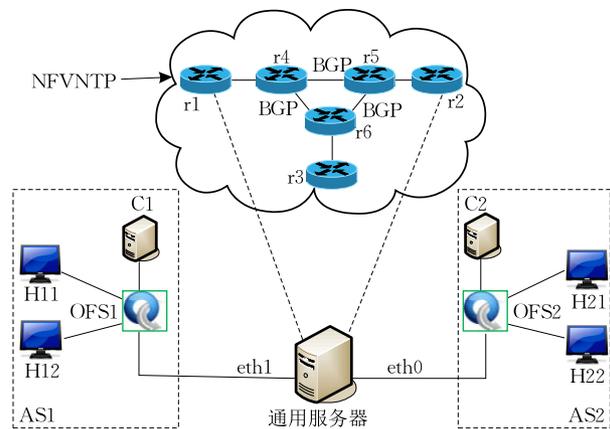


图 10 OpenFlow 网络通过广域网互连

为了测试这种混合网络,我们需要做如下工作:(1)在 SDN 控制器 C1 和 C2 上分别编写相关控制程序,为通信主机通向广域网路由器建立流表项;(2)在主机 H11 上配置 VLC 服务器,采用 HTTP 协议连续播放视频,而在主机 H12、H21 和 H22 上配置 VLC 客户端,分别访问 H11 上正在播放的视频;(3)在服务器主机安装 Wireshark,俘获视频流

的通信过程.图 11 为视频流随时间变化的曲线,这 3 条流量曲线分别表示同时有 1、2、3 个客户访问视频服务器时的视频流量随时间变化的情况.

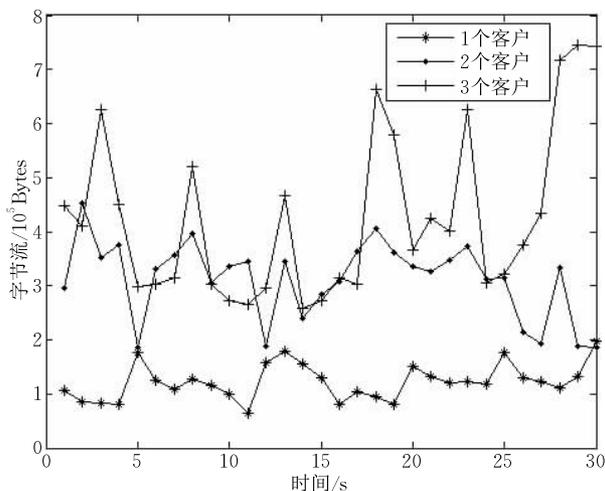


图 11 视频流变化曲线

试验 3. 结果表明:(1)这种具有 IP 和 SDN 混合结构的网络具有很好的互联互通性,无需对服务器端和客户端的软件进行任何修改;(2)通过在控制器上编写程序较容易控制视频流的行为,这种结构的网络具有很强的控制能力;(3)通过虚实互联互通,NFVNTP 能够很好地支持广域网控制功能.

就上面 3 个试验例子而言,如果建立真实的网络环境来做试验,就代价而言,对于试验 1 和试验 3 代价虽然可以接受,但也相当可观;而对于试验 2,代价高昂,无法接受.对于时间而言,试验 1 至少花费一周时间,试验 2 则至少花费 1 年时间,而试验 3 也需要数月时间.对于安全性而言,在真实环境中,进行试验 2 则可能导致数据链所使用的频率泄露.但是,这些试验在 NFVNTP 上进行时,它的成本、周期和安全性等方面的优势就能展现出来.

8 结束语

随着网络空间安全、网络智能、网络大数据、未来网络、网络管理等领域的飞速发展,对网络科学和技术的研究不断深入,在实际环境下用试验手段测试验证相关网络理论和技术的这个“最后一公里”问题,已经成为我们必须面对的难题.“工欲善其事,需先利其器”,我们亟需发展一种真实、高效、经济的新型试验平台.本文研究的基于 NFV 技术的网络试验平台 NFVNTP 具有测试保真性、易构建、成本低、安全保密、接口开放等优点,能够为网络科学和

工程技术研究提供先进的基础设施. 本文提出了网络试验平台测试逼真性的概念, 即要求平台提供测试结果的真实性、与网络环境的兼容性和应对多种应用场景的适应性. 基于网络功能虚拟化的优势, 本文提出了基于 NFV 的网络试验平台(NFVNTP)的体系结构和平台组件模型, 并给出 NFVNTP 的一种设计方法. 据我们所知, 这种试验平台技术及其设计方法为我们首次提出. 下一步工作包括, 根据当前网络研究和应用的实际需求进一步优化 NFVNTP 结构和性能, 扩展平台的规模和功能, 开发图形化操作界面提升平台的可用性等.

参 考 文 献

- [1] Rashid M, Joan S, Juan-Luis G, et al. Network function virtualization: State-of-the-art and research challenges. *IEEE Communications Surveys and Tutorials*, 2016, 18(1): 236-262
- [2] Yuan Xiao, Cai Zhi-Ping, Liu Shu-Hao, et al. Large scale network emulation software and its key technologies. *Computer Technology and Development*, 2014, 24(9): 9-12 (in Chinese)
(袁晓, 蔡志平, 刘书昊等. 大规模网络仿真软件及其仿真软件. *计算机技术与发展*, 2014, 24(9): 9-12)
- [3] Salopek D, Vasic V, Zec M, et al. A network testbed for commercial telecommunications product testing//*Proceedings of the 22nd International Conference on Software, Telecommunications and Computer Networks*. Split, Croatia, 2015: 372-377
- [4] Yan L, McKeown N. Learning networking by reproducing research results//*Proceedings of the ACM SIGCOMM Conference on SIGCOMM*. Los Angeles, USA, 2017: 19-26
- [5] Pridmore L, Lardieri P, Hollister R. National cyber range (NCR) automated test tools: Implications and application to network-centric support tools//*Proceedings of the*

AUTOTESTCON. Orlando, USA, 2010: 1-4

- [6] Soares J, Dias M, Carapinha J, et al. Cloud4NFV: A platform for Virtual Network Functions//*Proceedings of the IEEE Conference on Cloud NETWORKING*. Boston, USA, 2014: 288-293
- [7] Vilalta R, Mayoral A, Muñoz R, et al. The SDN/NFV cloud computing platform and transport network of the ADRENALINE testbed//*Proceedings of the IEEE Conference on Network Softwarization*. London, UK, 2015: 1-5
- [8] McKeown N, Anderson T, Balakrishnan H, et al. OpenFlow: Enabling innovation in campus networks. *ACM SIGCOMM Computer Communication Review*, 2008, 38(2): 69-74
- [9] Abeni L, Kiraly C, Li N, et al. On the performance of KVM-Based virtual routers. *Computer Communications*, 2015, 38(5): 40-53
- [10] Bugnion E, Devine S, Rosenblum M, et al. Bringing virtualization to the x86 architecture with the original vmware workstation. *ACM Transactions on Computer Systems*, 2012, 30(4): 1-51
- [11] Batalle J, Riera J F, Escalona E, et al. On the implementation of NFV over an OpenFlow infrastructure: routing function virtualization//*Proceedings of the IEEE Software Defined Networks for Future Networks and Services*. Trento, Italy, 2013: 1-6
- [12] Cerf V, Burleigh S, Hooke A, et al. Delay-tolerant networking architecture. Heise Zeitschriften Verlag, 2007
- [13] Scott K, Burleigh S. Bundle protocol specification. *Work in Progress Internet Draft*, 2007
- [14] Zuo Qing-Yun, Chen Ming, Zhao Guang-Song, et al. OpenFlow-based SDN technologies. *Journal of Software*, 2013, 24(5): 1078-1097 (in Chinese)
(左青云, 陈鸣, 赵广松等. 基于 OpenFlow 的 SDN 技术研究. *软件学报*, 2013, 24(5): 1078-1097)
- [15] Jain S, Kumar A, Mandal S, et al. B4: Experience with a globally-deployed software defined wan//*Proceedings of the ACM SIGCOMM Conference on SIGCOMM*. Hong Kong, China, 2013: 3-14



CHEN Ming, born in 1956, Ph. D., professor, Ph. D. supervisor. His main research interests include network functions virtualization, network performance analysis and modeling, UAV network, future networks.

TAO Xiao-Mei, born in 1991, M. S. candidate. Her main research interests include network functions virtualization, software defined networking.

HU Chao, born in 1984, Ph. D., lecturer. His main research interests include network architecture and distributed computing.

XU Bo, born in 1980, Ph. D., associate professor. His main research interests include network measurement and network applications.

XING Chang-You, born in 1982, Ph. D., associate professor. His main research interests include network measurement and future networks.

Background

Before new network equipment and protocols being entered a practical stage, the most critical process should be to test and validate them in the real network environment thoroughly. However, it may be very difficult to test in such way for the important infrastructure such as the space-air-ground network. To settle the problem, this paper proposed the concept of test fidelity, the architecture of network functions virtualization based network testing platform (NFVNTF) and its component model, and a method of

designing NFVNTF were given.

This work is supported by the National Natural Science Foundation of China (No. 61379149) and the National Key Basic Research and Development Program of China (No. 2012CB315806), which research measurement technology of future networks. During the period, we found there exist many shortcomings for current simulation methods; especially they are hard to guarantee the fidelity of simulating behaviors and performance of large scale of or complex networks.

《计算机学报》