

未来互联网试验平台研究综述

梁军学¹⁾, 林昭文¹⁾, 马严^{1),2)}

¹⁾(北京邮电大学 网络技术研究院信息网络中心 北京 100876)

²⁾(北京邮电大学 智能通信软件与多媒体北京重点实验室 北京 100876)

摘要 相对于互联网上层应用和底层通信技术的蓬勃繁荣, 互联网的基本网络架构并没有得到同步发展, 还是基于设计之初的 TCP/IP 协议。随着应用环境、用户需求和底层设施的不断发展变化, 现有的网络结构已暴露出越来越多的缺陷。针对这些问题, 研究人员提出了许多新的网络架构、概念、服务、协议和算法, 以构建不同于现有互联网的新型未来互联网。不过, 任何一种新的思想在实现或部署到真实的网络之前, 都必须经过严格的测试和验证。因此, 构建大规模的、真实的、可控的试验平台不仅是未来互联网研究的基础, 也是对未来互联网本身探索和研究的途径。本文从规模、功能、开放性等基本特性, 以及科学试验要素、网络元素类型三个方面对试验平台进行了归纳和分类; 分析了试验平台的结构化原则和分层结构; 重点分析和总结了构建试验平台的关键技术, 包括试验描述语言、控制框架和网络资源分片技术; 分析了国内试验平台的发展现状; 最后探讨了未来互联网试验平台的发展趋势。

关键词 未来互联网; 试验平台; 全球网络创新环境; 控制框架; 资源分片

中图法分类号 TP393 DOI 号:

Research of Future Internet Experiment Platform: A Survey

LIANG Jun-Xue¹⁾, LIN Zhao-Wen¹⁾, MA Yan^{1),2)}

¹⁾(Research Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing 100876)

²⁾(Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract Comparing to the booming prosperity of the Internet application and the underlying communication technology, the basic network architecture of the Internet has not been developed simultaneously, and is based on the beginning design of TCP/IP protocol. With the evolvement of application environment, user requirements and underlying facilities, the existing network architecture has exposed more and more defects. For these problems, researchers have proposed many new network architecture, concepts, services, protocols, and algorithms to build the new Future Internet, which is different from the existing Internet. However, any of such new ideas must be tested and evaluated rigorously before being realized or deployed in the actual Internet. So it is not only the base of Future Internet research, but also the way to explore the Future Internet itself to build the large-scale, realistic, controlled experiment platform. This paper concludes and categorizes the experiment platform from the basic characteristics of scale, function, openness, science experimentation elements, and network element type. The architecture principle and hierarchical framework are analysed. Apart from that, the key technologies to build Future Internet, including the experiment description language, control framework and network resource

收稿日期: 2012年5月22日; 最终修改稿收到日期: 2013年5月11日。本课题得到国家高技术研究发展计划(863计划)(No.2011AA010704, No.2013AA014702)、国家CNGI专项(CNGI-12-02-027)资助。非计算机学会会员。梁军学, 男, 1982年生, 博士研究生, E-mail: liangjx@buptnet.edu.cn, 主要研究领域为云计算、网络虚拟化。林昭文, 男, 1979年生, 博士, E-mail: linzw@buptnet.edu.cn, 副教授, 主要研究领域为未来互联网关键技术、网络安全。马严, 男, 1955年生, 学士, E-mail: mayan@bupt.edu.cn, 教授, 主要研究领域为未来互联网、IPv6和网络安全。

slicing, are analyzed mainly. After that, the status of domestic experiment platform is analyzed. Finally, the development trend of Future Internet experiment platform is discussed.

Key words Future Internet; Experiment Platform; GENI; Control Framework; Resource Slicing

1 引言

经过 40 多年的发展, 互联网已经成为与人们工作、生活息息相关、不可缺少的基础设施之一。然而, 相对于互联网上层应用和底层通信技术的蓬勃繁荣, 互联网的基本网络架构并没有得到同步发展, 还是基于设计之初的 TCP/IP 协议。随着应用环境、用户需求和底层设施的不断发展变化, 现有的网络结构已暴露出越来越多的缺陷。例如, 互联网的商业化带来了网络安全问题, 无线通信的发展引入了设备的移动性问题, 物联网的兴起又使低能耗成为关注点之一, 此外, IP 地址资源的枯竭, 服务质量的保证, 网络管理的部署, 路由的可扩展性问题, 都给现有互联网的发展带来巨大挑战。

针对上述问题和挑战, 研究人员提出了许多新的网络架构、概念、服务、协议和算法, 以构建不同于现有互联网的新型未来互联网。从研究思路来看, 这些研究可分为三种路线: 一种是“改良式”^[1], 即针对现有互联网体系结构存在的不足进行增量式修补; 第二种是“革命式”, 即构建一套全新的网络体系架构, 以从根本上解决互联网面临的问题^[1]; 第三种是介于两者之间的“演进式”路线^[2], 即保持互联网的核心和设计原则相对稳定, 谨慎地改变约束其扩展的基本要素。改良式路线具有较强的延续性与可操作性, 短期内能够取得较好的效果, 但是长期来看, 补丁式的改进反而可能会降低整个互联网架构的灵活性、可靠性和可管理性, 难以从根本上解决互联网所面临的问题。革命式路线虽然有好的愿景, 但需要重新构建网络, 在实际中难以大规模部署和验证, 提出的很多方案由于部署代价难以承受, 至今尚未得到大规模的实际应用。演进式路线试图在两种路线之间取得较好的折衷和平衡, 以在满足互联网需求不断扩展的同时又能实现稳定过渡, 但能否找到一个为各方所接受的平衡点并一致推动, 目前来看也面临着不小的困难。

无论采用何种路线, 任何一种新的算法、协议或体系架构在部署或应用前, 都必须经过严格的测试和验证, 都需要构建相应的试验验证平台。从互

联网的发展历史来看, 互联网本身就起源于试验网, 作为互联网雏形的阿帕网, 最初就是只有几个节点规模的试验网。因此, 未来互联网的发展也必须且只能通过试验平台的构建来支持和推动。从一定意义上说, 这些试验平台就是新型未来互联网的雏形, 对试验平台的研究本身就是对未来互联网探索的一种重要方式。

本文余下的内容组织如下: 第 2 节归纳和总结了试验平台的分类; 第 3 节分析了试验平台的体系架构; 第 4 节至第 6 节分别分析和总结了构建试验平台的关键技术, 包括试验描述语言、控制框架和网络资源分片技术; 第 7 节分析了国内试验平台的发展现状; 第 8 节展望了未来互联网试验平台的发展趋势。

2 试验平台分类

未来互联网试验平台, 顾名思义, 就是为未来互联网的研究、测试、验证而提供的计算机硬件和软件操作环境。关于试验平台的分类, 目前还没有统一的标准, 也未见相关文献分析。一般而言, 我们可根据试验平台的功能、规模和开放性等基本特性进行分类, 但这种分类无法反映试验平台的技术特点和对未来互联网研究的支持作用。因此, 除按照上述特性对试验平台进行基本分类外, 本文还从试验要素和网络元素两个方面对试验平台进行了归纳和分类。

2.1 基本分类

根据试验平台的部署规模划分, 可分为全球级、国家/地区级、本地级三类。全球级试验平台如 PlanetLab 等是分布于世界各地的、具有全球影响力的大规模试验平台, 是各个国家和地区间开展网络研究合作的重要依托。国家/地区级试验平台一般由国家/地区的政府、高校和相关科研机构发起和建设的较大规模试验平台, 如美国的 GENI (Global Environment for Networking Innovations)、欧盟的 FIRE (Future Internet Research and Experimentation)、日本的 JGN-X (Japan new

Generation Network testbed)¹、韩国的 K-GENI (Korea GENI)^[3] 以及我国的 CNGI (China Next Generation Internet) 等。本地级试验平台通常由某一高校或科研机构独立建设，其规模受限于本地站点的资源数量，典型的有美国华盛顿大学的开放网络试验室 (ONL, Open Network Laboratory)^[4]、美国海军研究试验室开发的 CORE (Common Open Research Emulator)²、意大利那不勒斯费德里克二世大学开发的 Neptune³ 等平台。

根据试验平台的功能划分，可分为专用型和通用型两类。专用型试验平台是指为了某种特殊网络试验而构建的平台。以欧盟的 FIRE 项目为例，其中的 BonFIRE⁴ 主要为大规模的云应用、服务和系统提供测试环境；SmartSantander⁵ 的主要目标是为智慧城市相关的应用和服务提供城域网规模的试验设施；EXPERIMEDIA⁶ 则主要为未来媒体网络 (FMI, Future Media Internet)^[5] 提供大规模的研究设施。通用型试验平台一般提供可由用户定制的虚拟网络环境，并提供相应的编程接口，允许试验人员上传其开发的试验软件。当然，通用型试验平台并非支持所有类型的试验，所提供的试验环境也受其底层采用的技术限制。例如，由于 PlanetLab 的节点是重叠在现有的 IP 网络之上，因此任何网络层的试验都会受到底层真实网络的影响而难以控制。而 Emulab、NetKit^[6]、Modelnet^[7] 等平台则采用软件来仿真网络设备的功能，难以保证较高的真实性和性能。

根据试验平台的开放性划分，可分为完全开放、有限共享和私有平台三类。完全开放是指试验平台开放给所有的试验人员，如美国的 PlanetLab、Emulab、ONL、欧盟的 OFELIA (OpenFlow in Europe Linking Infrastructure and Applications)⁷ 等，试验人员只需注册为成员并提供试验目的等说明，就可以申请试验资源并完成试验。有限共享是指试验资源仅在特定的组织或机构间共享，如欧盟的 OneLab⁸ 允许欧洲范围内的科研机构使用，我国的 DragonLab^[8] 则在会员单位之间共享试验资源。私有平台则仅限于建设试验平台的组织或机构内部使用，如以上提到的 CORE、Neptune、NetKit、

Modelnet 等试验平台均为私有平台。

2.2 根据试验要素分类

试验平台首先是一种科学试验工具，因此可以根据科学试验中的三要素，即试验者、试验对象、试验手段来分析。

试验者是试验活动和认识的主体，不同试验者由于拥有的资源、服务的对象、研究的主体不同，其对试验平台或工具的要求也不同。未来互联网的推动力量来自两个阵营，一是网络服务的提供者，即电信运营商，另一方是网络服务的使用者，即互联网用户。一般而言，电信运营商所构建的试验平台多基于真实的物理设备和用户流量，功能上侧重于对传输等服务的优化及网络的管理，如欧盟的 PII (Panlab Infrastructure Implementation)⁹ 项目，其主要目标就是开发高效的技术和机制来实现欧洲现有电信试验网的联合，为验证新的运营级服务、网络技术和商业模型提供真实的测试环境。互联网用户或社区所构建的试验平台多基于虚拟化和仿真技术，功能上侧重于新应用、协议和算法的验证，如欧盟的 FEDERICA (Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures)^[9] 项目，其主要目标就是将分布在欧洲范围内的计算机和网络试验资源联合起来，构建一个端到端隔离的试验环境，用于验证网络安全、分布式协议和应用、以及 clean-slate 式的未来互联网体系结构。

试验对象是试验者所要研究的对象，是试验活动和认识的客体。未来互联网研究中的试验对象就是研究人员为了改进或变革网络所提出的新的网络架构、概念、服务、协议和算法等。以未来互联网体系架构为例，目前学术界已围绕互联网的可扩展性、动态性、安全可控性等问题提出了多种解决方案^[10]，如以 CCN (Content-Centric Network)^[11] /NDN (Named Data Network)^[12] 为代表的面向可扩展性的体系结构，以 MobilityFirst^[13] 为代表的面向动态性的体系结构，以 SOFIA (Service Oriented Future Internet Architecture)^[14] 为代表的面向服务体系结构，以及国内张宏科教授等提出的两层可信网络体系结构^[15]、汪斌强教授等提出的可重构柔性网络架构^[16] 等。为了研究这些体系结构的可行性，都需要有相应的试验平台来支持。

试验手段是试验者和试验对象之间的中介，由

1 JGN-X. <http://www.jgn.nict.go.jp/>

2 CORE. <http://cs.itd.nrl.navy.mil/work/core/>

3 Neptune. <http://code.google.com/p/neptune-network-emulator/>

4 BonFIRE. <http://www.bonfire-project.eu/>

5 SmartSantander. <http://www.smartsantander.eu/>

6 EXPERIMEDIA. <http://www.experimedia.eu/>

7 OFELIA. <http://www.fp7-ofelia.eu/>

8 OneLab. <http://www.onelab.eu/>

9 Panlab. <http://www.panlab.net/>

试验仪器、工具、设备等客观物质条件组成,在一定意义上就是试验平台本身。根据不同的试验手段和验证方法,可将试验平台分为模拟软件、仿真平台、试验床和试验网络四种。模拟软件(NS2、NS3等)是一种面向对象的网络模拟器,可根据试验的网络拓扑和特性建立相应的数学模型,并允许用户通过试验脚本来控制试验。模拟软件虽具有控制灵活、成本低等优点,但其最大的缺点是缺乏真实性,无法反映真实网络的状态和影响。仿真平台如Emulab、CORE等采用虚拟化的服务器集群来仿真分布式的网络,在一定程度上解决了真实性问题,但仿真平台一般采用基于软件的方式来仿真网络设备,其性能难以与真实硬件相比,另外仿真平台多数是单个站点内的封闭系统,因此部署规模受到限制,难以支持大规模的网络试验。试验床一般通过重叠网等技术将分布在多个站点的试验网络连接起来,并通过真实的互联网来传输试验流量,因此在真实性和规模上都有较好的保证。但对试验的部署、控制以及回放缺少有效的支持。典型的试验床有PlanetLab、OneLab等。试验网络是指具有试验性质的生产网络,如美国的Internet2和我国的CERNET2(China Education and Research Network)等,它们既作为与商业运行网络类似的基础设施,同时又承担着科学研究项目的试验和验证工作,为保证网络的正常运行,试验网络所承载的试验都有严格的要求,一般不承载具有破坏性创新的试验。

2.3 根据网络元素分类

试验平台包括计算、存储、网络等资源,其中网络资源的特性决定了试验平台功能和性能,因此可根据组成网络资源的网络元素(设备)来分析。试验平台中的网络元素可分为软件和硬件两类,其中软件包括模拟器和仿真器,硬件包括商业网络设备、基于NetFPGA(Field Programmable Gate Array)的网络设备、基于网络处理器的网络设备以及新型的OpenFlow¹交换机等。

网络模拟器是一种脱离真实网络而独立运行的软件程序,它通过建立实际网络系统的数学模型并按照相同的运行机理来模拟真实系统的动态行为。例如NS2模拟器中的网络元素就是一个面向对象的类实例,通过相应的成员变量和方法来提供网络状态的控制以及与其它网络元素的交互。网络仿真器使用仿真接口作为真实网络应用程序和仿真

对象之间的桥梁,它可以接收与被仿真对象一样的数据,执行同样的程序并获得相应的结果。与模拟器不同的是,仿真器与真实网络之间存在交互,它通过软件的方式为网络协议或应用构建一个虚拟的网络运行环境并影响真实的数据流。如Emulab、CORE、Neptune等仿真平台就采用软件交换机和TC/Netem、Dummynet^[17]等链路仿真工具来仿真网络设备和链路。

商业网络设备是指由制造商提供的用于生产网络的硬件设备,与模拟器和仿真器相比,它具有较高的包处理性能和更好的规模扩展性,但现有的网络设备都是封闭的系统,仅提供有限的功能和已知的服务,只能转发标准的或预定义的数据包,不具有动态性和灵活性。为支持研究人员对已有的数据包进行修改或引入自定义的数据包,需要网络设备提供可编程的网络接口。典型的可编程网络设备有NetFPGA和网络处理器(Network Processor, NP),如ONL就使用了基于网络处理器的路由器作为用户可定制的IP处理设备。

OpenFlow(OF)交换机是一种新型的可编程网络设备,其基本原理是将网络设备的控制面从硬件中独立出来,交由一个集中式的软件控制器来决定每个数据包的流向,从而实现了数据转发和路由控制的分离^[18]。表1总结了不同类型网络元素的性能和灵活性等特性,从表中可以看出,OF交换机具有与商业网络设备同等的性能、易扩展性以及实现的低复杂性,同时又具有开放的、灵活的可编程能力,因而逐渐成为构建未来互联网试验平台的首要技术^[19],许多试验平台如GENI、OFELIA、OpenLab²等都采用了OpenFlow技术。

表1 不同类型网络元素特性比较

	性能	扩展性	灵活性	复杂性	真实流量	开放
模拟器	中	中	好	中	否	是
仿真器	低	差	好	中	是	是
商业网络设备	高	好	差	低	是	否
NetFPGA	高	差	中	高	是	是
网络处理器	高	中	中	高	是	是
OF交换机	高	好	好	低	是	是

3 试验平台体系架构

从以上对试验平台的分类研究中可以看出,试

1 <http://www.openflow.org/>

2 OpenLab. <http://www.ict-openlab.eu/>

验平台不仅在规模、功能和开放性上具有多样性特点，而且由于参与的试验要素和采用的网络元素不同而呈现出复杂多变的特点。为更好地认识这种多样性和复杂性，需要理解隐藏在各种平台背后的体系架构。根据 SunTone 体系架构方法学的定义，“体系架构是一组结构化原则，这些原则使得系统能够由一些较为简单的系统组成，它们有彼此独立的环境，但作为一个整体的大系统时又能彼此保持一致。”因此，分析试验平台的体系架构，就是根据其设计目标或结构化原则，将试验平台划分为若干独立的层次或模块，通过分析这些层次或模块之间的交互机制，进而理解和把握构建试验平台的关键技术。

3.1 试验平台的结构化原则

根据对 GENI、FIRE 等项目系统需求和建设目标的分析，试验平台的结构化原则可归纳为以下五个：

1)统一性：指试验平台应为用户提供统一的试验抽象模型，以允许用户定义符合标准的、可重用的试验描述。

2)可分片性：指试验平台的资源可通过虚拟化技术逻辑切片，以允许多个试验共享资源。

3)隔离性：指用户的试验之间互不影响，若试验平台构建于生产网络之上，还指用户的试验不会影响生产网络。

4)可编程性：指提供给用户的资源应具有可编程能力，以允许用户开发和部署新的架构、应用、协议和算法。

5)扩展性：指试验平台的规模和资源类型是易于扩展的，以满足试验对大规模的、真实异构的资源需求。

3.2 试验平台的分层结构

试验平台本质上是构建于现有互联网上的一种服务平台，因此其分层结构可以按照互联网的用户、服务提供商、基础设施提供商三层模型划分为试验表示层、试验服务层和试验设施层。其中试验表示层是用户试验的抽象模型，它通过一组规范化的语义抽象描述了一个试验的本质属性和生命周期，为试验服务层提供了服务的对象实体。试验服务层是试验服务的提供者，它利用基础设施提供的试验资源，通过动态的调度组合实现资源的优化分配，并将抽象的试验对象实体映射到相应的资源分片上，同时通过一定的机制实现试验的隔离、控制、

监测和管理等功能。试验设施层是试验服务的承载者，它通过虚拟化等技术，屏蔽底层物理资源的异构性，为试验服务提供抽象的资源池和访问资源的统一接口。

3.3 试验平台的关键技术

根据上述结构化原则和体系架构描述，在设计未来互联网试验平台时，需着重解决如下三个关键问题：

1)试验描述语言，它提供了一组规范化的语义来描述网络试验的网络拓扑和试验的部署、控制、输出等特性，为用户管理试验的生命周期提供统一的视图。试验描述语言是用户和试验平台之间交互的接口，它解决的是试验平台的统一性问题。

2)控制框架，它定义了一组管理实体和实体间的通信接口及流程，负责调配、组装试验平台的资源和服务，以为用户提供隔离的、可编程的、可扩展的试验环境。控制框架是试验平台的核心，它解决的是试验平台的隔离性、可编程性和扩展性问题。

3)网络资源分片，即底层的试验网络可同时被多个用户共享，且每个用户都认为独占所分配的资源分片。网络资源分片是试验平台的基础，它解决的是试验平台的可分片性问题。

4 试验描述语言

作为试验平台的重要组成部分，试验描述语言的研究目前还没有受到足够的重视，除传统的网络模拟软件具有较完善的试验描述语言外，大多数试验平台的试验描述语言还在发展的过程中。本节首先对试验描述语言进行分类，然后按照分类逐一进行分析。

4.1 试验描述分类

如 2.2 节所述，试验平台根据其所采用的试验手段不同，可分为模拟软件、仿真平台、试验床和试验网络四种。模拟软件一般采用标准化的编程语言或脚本语言来描述试验，并通过模拟器来解释和执行试验过程。这种试验描述语言与模拟器的执行具有很高的耦合度，可完整地定义和描述试验中的每个元素和过程，本文称之为面向过程的试验描述。试验床在本质上是一种分布式网络平台，往往分布于不同的研究机构，分属于不同的管理域，因而难以实现自动的试验软件部署和试验过程的控

制管理,目前大多数试验床仅提供对试验资源的描述,本文称之为面向资源的试验描述。仿真平台对试验的控制和管理能力介于模拟平台和试验床之间,因此,其试验描述与仿真平台内在的实现机制紧密相关,除可采用面向资源和面向过程的试验描述之外,还可采用一种混合的描述方法,如在资源描述的基础上增加试验软件的部署等描述,但试验的控制或结果输出等仍然交由用户手动完成,本文称之为面向平台的试验描述。试验网络由于是一种生产网络,因此其试验通常不需要描述,而是直接部署到网络中。下面分别分析三种不同的试验描述。

4.2 面向过程的试验描述

OMNeT++ (Objective Modular Network Testbed in C++)¹是一种面向对象的离散事件模拟工具,在模型描述、网络拓扑定义、模型实现、跟踪支持、调试等方面都有较强的优势。OMNeT++采用NED (Network Description)和C++两种语言来建模。其中,NED用来描述模型的拓扑结构,NED虽然简单,但具有丰富的拓扑定义功能,且可以实现动态加载,便于更新仿真模型的拓扑结构;C++则用来实现模型的构建和消息的处理等功能。SSFNet²是一个Internet网络及其协议的建模软件,由基于Java SSF(Scalable Simulation Framework)的组件构成,主要支持IP包级别上的细粒度模拟。SSFNet使用领域建模语言(DML, Domain Model Language)程序建模网络。DML程序利用简化过的语言表示各种网络设备、连接和协议,而不是强制网络建模者使用SSF实现语言(通常是C++或Java)进行编程。与标准的编程语言相比,DML具有更强的结构性,并对网络模拟制定了标准的格式,因而具有更强的可读性。其它的模拟软件,如NS2、NS3、J-SIM³,GloMoSim⁴,OPNET⁵,QualNet⁶,在网络建模的能力上各有区别,但就其建模语言或试验描述而言,都采用了基于高级编程语言或脚本语言的方式,因此均属于面向过程的试验描述。

4.3 面向资源的试验描述

以GENI为代表的试验床多数采用了面向资源的试验描述。这里的资源既可以是物理资源,如路

由器和交换机,也可以是逻辑资源,如CPU时间或无线频段。从用途来看,资源描述可用于站点的资源发布,用户的资源请求和平台的资源分配三方面。可见,资源描述是用户、基础设施提供者和试验服务提供者三者之间的黏合剂,是试验平台中各实体间交互的基础。RDF (Resource Description Framework)⁷是W3C (World Wide Web Consortium)在XML (Extensible Markup Language)的基础上推荐的一种标准,用于表示任意的资源信息。NDL (Network Description Language)⁸是在RDF的基础上发展起来的一种用于描述物理网络的描述语言。NDL定义了一组实体和属性,为描述复杂的网络拓扑提供了一组共享的通用语义,被广泛应用于需要描述网络的项目中。例如,GENI的子项目集ORCA就采用了NDL-OWL⁹来描述网络资源,NDL-OWL利用OWL (Ontology Web Language)对NDL作了进一步扩展,增加了新的类层以描述服务器、虚拟机、云服务、存储服务器等边缘资源。为了在各子项目间提供资源共享和互操作服务,GENI定义了统一的资源规范RSPEC (Resource Specification)¹⁰,不仅对异构资源(如计算机网络、无线网络、传感器网络和存储网络等)的描述作了统一的规范,也将测量等带有资源性质的服务纳入到描述的框架中。与GENI类似,FIRE子项目PII也在其Teagle¹¹架构中利用RADL (Resource Adapter Description Language)^[20]来描述资源适配器(RA, Resource Adapter)中间件,其中RA是用户和试验床资源间的代理,用于资源的请求、发布和分配。

4.4 面向平台的试验描述

资源描述是试验描述的基础,但仅有资源描述对于试验平台而言还远远不够。例如,对于大规模的网络试验,如果试验软件的部署和软件命令控制均由人工完成,则试验的效率和效果将大打折扣。Emulab采用了以Tcl编写的扩展NS脚本来配置试验,但由于NS中的实体与Emulab中的虚拟资源具有复杂的关系,因此在句法层次上难以准确解析。OEDL (OMF Experiment Description Language)^[21]是OMF控制框架的试验描述语言。它是一种特定领域的叙述性语言,它除了可以描述需要的资源以及配置和连接方式外,还可以定义试验本身的协调和相

1 OMNeT++. <http://www.omnetpp.org/>

2 SSFNet. <http://www.ssfnet.org/>

3 J-Sim. <http://www.j-sim.zcu.cz/>

4 GloMoSim. <http://pcl.cs.ucla.edu/projects/glomosim/>

5 OPNET. <http://www.opnet.com/>

6 Qualnet. <http://www.scalable-networks.com/content/>

7 RDF: Resource Description Framework. <http://www.w3.org/RDF/>

8 NDL: Network Description Language, <http://sne.science.uva.nl/ndl/>

9 NDL-OWL. <https://geni-orca.renci.org/trac/wiki/NDL-OWL/>

10 GENI RSpec. <http://groups.geni.net/geni/wiki/GeniRSpec/>

11 <http://www.fire-teagle.org/>

关上下文。OMF的试验描述以Ruby语言编写，可分为资源描述/配置和事件/任务描述两部分，OMF的试验控制器在试验运行期间会监视定义的事件，并在事件发生时执行相关的任务。NEPI (Network Experimentation Programming Interface)^[22]的目标与OEDL类似，它是一个网络试验生命周期的管理工具，其长远目标是为任意试验平台提供一致的接口，使研究人员可以使用统一的工具就可以在模拟平台、仿真平台和试验床甚至三种的混合平台上完成试验。NEPI使用python脚本来描述试验，目前已支持NS3 模拟平台，Netns仿真平台¹和PlanetLab试验床。

5 试验平台控制框架

控制框架的主要任务是管理底层的资源和服务，由于试验平台的资源可能分布在不同的站点，所提供的服务可能采用不同的技术，同时，多个试验平台之间也需要共享资源和服务，因此，控制框架除管理本地资源和服务外，还需要管理跨站点跨平台的资源和服务。本节首先介绍 GENI 的通用控制框架，然后以 PlanetLab 为例，介绍多平台的联邦框架，最后介绍 PanLab 的基于 Web 的联邦框架及其演进。

5.1 GENI通用控制框架

GENI 的控制框架可分为两种，一种是通用控制框架，规定了控制框架实现的一般架构；另一种是专用控制框架，是特定试验平台控制框架的具体实现。本文仅介绍通用控制框架，专用的控制框架实现可参阅相关文献^{[23][24]}。GENI 通用控制框架主要包括组件管理器、集合管理器、服务管理器和结算中心。其中组件管理器实现了组件内的资源虚拟化和资源分隔，并提供可编程能力和安全保证；集合管理器负责集合内的资源管理（如授权、资源调度分配）以及集合间的资源共享；服务管理器则为用户试验提供特定的服务支持；结算中心是控制框架的核心，它是一个管理和注册中心，负责维护用户、资源切片和组件的信息，这些信息可用于组件或集合的访问控制、控制策略、信任机制和联邦机制，另外还包含了可选的票据日志和软件仓库，以保存资源使用记录和管理 GENI 提供的软件。此外，控制框架还定义了所有实体之间的接口，消息类

型，试验过程中实体之间的消息流以及在实体之间传送消息的控制接口。

5.2 多平台联邦框架

PlanetLab 的联邦架构可分为以下四种，其中前两种是同构平台间的联邦，后两种是异构平台间的联邦。

多集合联邦：多个平台共享同一个注册表，PlanetLab 的分片管理器从注册表中获取各个平台上的资源授权，然后利用 PlanetLab 的集合管理器来创建分片。

完全联邦：每个试验平台都实现了分片管理和注册表功能，因此每个平台都可独立于其它平台。使用某个试验平台的用户可以在另一个试验平台上的组件内创建和管理资源切片。

基于替代分片管理器：其它试验平台的用户通过自身的分片管理器访问 PlanetLab 的注册表以获得资源授权，然后通过 PlanetLab 的集合管理器和组件管理器交互来完成分片的创建和管理。

基于通用注册表：PlanetLab 和其它试验平台共享同一个通用注册表，所有资源分片的授权仍由 PlanetLab 的注册表集中管理，但其它试验平台可直接调用自己的集合管理器来完成分片的创建和管理。

5.3 基于 Web 的联邦框架及演进

PanLab 各站点间的联邦采用了基于 web 服务的方式，用户可以通过 web 来查询和请求可用的资源。可用的资源通过元数据的形式保存在数据库中，PanLab 需要提供处理机制来标识、定位和分配多个站点的资源。为平滑地将多个异构的试验平台联邦为一个统一的平台，PanLab 采用了逐步演进的方法，以逐步提高资源发现和分配的自动化水平，PanLab 演进分为以下三个阶段。

集中式：这是演进的第一阶段，每个参与的站点在 Web 门户上手动公布其试验床的描述和共享资源的细节。目前 PanLab 已实现了基于 Web 的搜索服务 Teagle，试验用户向 Teagle 提交试验资源的请求，由 Teagle 在试验床元数据库中查找匹配的资源。

手动配置：在这个阶段中，参与的站点利用特殊的中间件发布其试验床元数据，并对外提供一个基础设施服务接口。Teagle 查询元数据库并通过该服务接口请求所需的资源，由于采用了虚拟化技术，资源的真实位置将会被隐藏。

¹ Netns emulator. <http://yans.pl.sophia.inria.fr/trac/nepi/wiki/netns>

按需配置：在此最后阶段中，Teagle 将根据用户的资源请求直接与虚拟资源层交互以自动地分配所需资源。Teagle 将提供尽力而为的配置服务，但复杂的配置可能仍需要用户直接访问分配的资源来完成。

6 网络资源分片

资源分片主要是以虚拟化技术为基础，目前服务器虚拟化和存储虚拟化已相对成熟，而网络虚拟化还有许多问题需要解决。本文主要讨论网络虚拟化中的网络资源分片机制，包括以PL-VINI¹为代表的基于重叠网的分片机制和以FlowVisor^[25]为代表的基于代理的分片机制。

6.1 可分片网络资源

为了虚拟化一个网络，首先要确定哪些资源需要被虚拟化或分片，Sherwood 等人^[25]认为主要有以下五种资源：

带宽：目前有几种较常见的方法来提供带宽的分割，如光网络中的 WDM，分组交换网络中的链路仿真等。

拓扑：每个分片应拥有自己的网络节点以及链路视图，能“感受”虚拟网络分片上的网络事件（例如链路失效和转发循环）。

业务流：广义的业务流既可以从/到某个地址集的所有分组，也可以是所有属于某个用户组的所有业务，或者特定的分组流，如所有的 http 流量、所有具有奇数以太网地址的流量等。

设备 CPU：交换机和路由器的计算资源也需要被分片，如果没有适当的 CPU 分片，交换机将停止转发低速路径上的分组和更新统计计数器，更重要的是，将停止处理转发表的更新（如路由变动）。

转发表：网络设备一般都支持固定数目的转发规则，转发表的分组应保证多个分片的包转发不会互相影响。

6.2 基于重叠网的分片机制

PL-VINI是在PlanetLab基础上构建的一个虚拟化服务层。PlanetLab的节点采用了linux VServer虚拟化技术，VServer提供了基于切片的资源控制，如CPU调度和带宽分割。PlanetLab的CPU调度器保证了每个切片能够“公平共享”节点的可用CPU资源，并支持临时的共享增长机制。PlanetLab中的网络分

割由Vnet模块来提供，Vnet记录并复用进出的业务流量，为每个分片提供对底层网络设备根级别的访问视图，每个分片都仅能访问属于自己的流量和特定的网络端口。PL-VINI的虚拟网络设备采用了UML (User Mode Linux)²，在每个用户的虚拟网络切片中，PL-VINI为网络拓扑中的每条链路在端点的UML实例中创建了一个隧道，因此运行在UML中的路由软件就可以在重叠的网络上拥有一个虚拟的网络。PL-VINI还修改了Linux的TUN/TAP设备内核以允许网络试验分片中的应用能够在重叠的虚拟网络上收发分组。

6.3 基于代理的分片机制

FlowVisor 是一个特殊的 OpenFlow 控制器，可看作是 OpenFlow 交换机和多个 OpenFlow 控制器间的透明代理。FlowVisor 通过流空间 (flowspace) 来分片业务流量。所谓流空间就是属于某个分片的一组 OpenFlow 流表项。FlowVisor 通过确保两个分片的流空间在拓扑中不重叠来分割两个分片。例如，两个试验人员分别通过控制器 A 和 B 来控制它们的 OpenFlow 网络分片，B 中发送到交换机的消息被 FlowVisor 截获，然后 FlowVisor 根据预定义的资源分配策略对消息进行透明地重写并发送到分片 B 所对应的交换机端口。与此类似，从交换机到客户控制器的消息也要先通过 FlowVisor 匹配，然后再转发到匹配的控制器的。FlowVisor 对带宽和 CPU 的分片依赖于具体的硬件或软件支持，如使用 VLAN 标签中的优先级比特来实现带宽分割，使用速率限制器来实现交换机的 CPU 分割等，但这些分割机制目前还并不完善，需要在未来工作中做进一步研究。

7 国内试验平台发展现状

我国政府和学术界也十分重视未来互联网及其试验平台的建设，如《“十二五”产业技术创新规划》明确将互联网新技术作为信息产业发展的重点，并早在 2003 年就启动了以CNGI (China Next Generation Internet)³为代表的下一代互联网项目。在平台建设方面，目前主要取得以下成果：一是以CNGI为依托，搭建了以IPv6 为核心的新一代互联网试验平台CERNET2，目前已成为世界上最大的纯IPv6 网络。二是由国家宽带网络与应用工程技术研

1 <http://www.vini-veritas.net/>

2 UML: User Mode Linux. <http://usermodelinux.org/>

3 CNGI. <http://www.cngi.cn/>

究中心等单位建设的下一代广播电视网（NGN, Next Generation Broadcasting network）国家试验床，重点支持柔性网络技术、三网融合技术、NGN技术、新型业务和商业模式的试验。目前已在上海等长三角城市展开规模部署，可支持网络电视、视频点播等大规模并发流媒体业务。三是由清华大学和锐捷网络等发起、国内多所高校共建的DragonLab项目，它采用了基于联邦架构的资源开发模型，重点集中于对流量研究、主动测量和BGP路由研究的支持，并提供了专用的客户端，以支持用户在远程以可视化的方式定制试验环境。

总体来看，与国外发达国家相比，我国在未来互联网试验平台建设上还存在一定差距。主要表现在以下几个方面：一是先进性和前瞻性不足，所建试验网络大部分主要是为了解决眼前和短期的问题，无论是 CERNET2 还是 NGN 都是基于原有的 IP 架构，缺乏对后 IP 时代网络的考虑。二是规模偏小，影响力不足，除 CERNET2 和 NGN 等具有商业特性的试验网络外，其它的试验平台无论是在资源数量还是连接站点数量都无法与国外 PlanetLab 等项目相比。三是缺乏与国外的合作和交流，除 IPv6 试验网与国外有部分合作外，很少有跨国的试验平台项目。

近期，国家高技术研究发展计划(863 计划)启动了“未来网络体系结构和创新环境”项目。该项目由清华大学牵头，汇集中科院计算所、北京邮电大学、东南大学、北京大学、工信部电信研究院、解放军信息工程大学、国防科大等国内重点科研院校和中兴、华为、华三、神州数码、锐捷等国内知名网络设备厂商，目标为构建支持未来网络体系结构和新协议创新的网络平台，即未来网络体系结构创新环境 FINE (Future network INnovation Environment)，为网络技术创新提供支撑作用。可以预见，FINE 的建设将进一步缩小我国与发达国家在未来互联网试验平台上的差距，加快我国未来互联网的技术创新和应用，提高我国在未来互联网乃至整个信息技术领域的核心竞争力。

8 总结

未来互联网演进的复杂性、不可确定性，决定了试验平台的建设必然是一个长期的、复杂的、逐步演进的过程。现有试验平台在资源类型、平台接口、试验服务等方面所呈现的多样性也反映出未来

互联网试验平台还处于动态的探索阶段，需要及时分析总结，以理解现有试验平台的体系架构和关键技术，把握未来互联网及其试验平台的发展趋势。通过以上的调研和分析，结合当前云计算、Web 服务以及软件定义网络等技术的发展，本文认为在建设我国未来互联网试验平台的过程中，需要把握以下三个趋势：

一是 Web 服务尤其是语义 Web 理论和技术的发展为研究开发统一的网络试验描述语言提供了可能。目前的试验平台采用了不同的试验描述语言，不同平台之间的试验描述脚本很难共享，使得大量的研究成果难以被同行验证或采用，不利于形成开放、透明、高效地研究体制。正因为此，未来互联网试验床的构建者已开始探索统一的试验描述语言，如 GENI 项目的 RSpec 首先从试验资源的层面进行统一，OEDL 和 NEPI 在试验的部署和控制方面进行了初步探索，NEPI 还试图对三种类型的平台语言进行统一，目前也取得了一定的进展。

二是以 OpenFlow 为代表的软件定义网络技术 (SDN, Software Defined Network)^[26]为未来互联网试验平台的构建提供了新的思路和方法。现有网络设备的封闭性使得网络技术的发展和创新的受到很大的局限，SDN 旨在实现网络互联和网络行为的可编程和开放式接口，从而支持未来各种新型网络体系结构和新型业务的创新。OpenFlow 技术作为 SDN 转发抽象的实现之一，已经受到学术界和工业界的普遍关注和广泛研究。目前包括 GENI、FIRE 计划中的很多项目已开始采用 SDN/OpenFlow 来构建试验平台，我国的 FINE 项目也明确采用 SDN 作为基础，并将实现具有中国自主知识产权的软件定义网络技术作为研究目标之一。

三是云计算尤其是基础设施云将成为构建试验平台基础设施的重要途径。目前的试验平台大多采用了专有的资源管理机制，给各个平台间的融合带来了很大的困难。近年来，Eucalyptus¹、OpenStack²、OpenNebula³、CloudStack⁴等开源云平台获得了蓬勃发展，它们对异构虚拟资源管理的能力为开发各种云计算平台提供了良好的基础设施服务。使用开源的云平台来构建试验平台可以使试验服务的提供与基础设施的提供相分离，基础设施的服务由开源的IaaS软件来提供，而试验服务只是

1 Eucalyptus. <http://www.eucalyptus.com/>

2 OpenStack. <http://openstack.org/>

3 CloudStack. <http://cloudstack.org/>

4 OpenNebula. <http://opennebula.org/>

做为其上的一种应用。这种架构使得试验平台很容易集成各种网络资源和第三方软件,因而具有更好的扩展性。如美国的ExoGENI^[27]项目就采用OpenStack开放云平台作为构建试验平台的基础。

致谢 在本文撰写和修改过程中,审稿专家和编辑老师提出了许多宝贵意见,在此表示衷心感谢。

参考文献

- [1] Rexford J, Dovrolis C. Future Internet architecture: Clean-slate versus evolutionary research. *Communications of the ACM*, 2010, 53(9): 36-40
- [2] Wu Jian-Ping, Lin Song, Xu Ke, Liu Ying, Zhu Min. Advances in evolvable new generation internet architecture. *Chinese Journal of Computers*, 2012, 35 (6): 1094-1108 (in Chinese)
(吴建平, 林嵩, 徐恪, 刘莹, 朱敏. 可演进的新一代互联网体系结构研究进展. *计算机学报*, 2012, 35 (6): 1094-1108)
- [3] K. JOOBUM, K. Dongkyun, Deployment of KREONET-GENI Future Internet Testbed, *IJACT: International Journal of Advancements in Computing Technology*, Vol. 3, No. 7, pp. 310 ~ 316, 2011
- [4] C. Wiseman, K. Wong, T. Wolf, and S. Gorinsky, Operational experience with a virtual networking laboratory. In *Proceedings of SIGCSE*. 2008, 427-431.
- [5] T. Zahariadis, G. Pau, G. Camarilo, *Future Media Internet*, *Communications Magazine, IEEE*, vol.49, no.3, March 2011, pp.110-111
- [6] M. Pizzonia and M. Rimondini, Netkit: easy emulation of complex networks on inexpensive hardware. In *Proceedings of TRIDENTCOM*. 2008
- [7] K.V. Vishwanath, A. Vahdat, K. Yocum, and D. Gupta, ModelNet: Towards a Data Center Emulation Environment. In *Proceedings of Peer-to-Peer Computing*. 2009, 81-82.
- [8] Wang Ji-long, Li Zhong-hui, Lv Guo-han, Jiang Cai-ping, Li Xing, Zhang Qian-li, DRAGON-Lab – Next generation Internet technology experiment platform. *Science in China Series F: Information Sciences*, 2008, 38(10): 1783-1794
(王继龙, 李钟辉, 吕国晗, 姜彩萍, 李星, 张千里. DRAGON-Lab—新一代互联网技术综合试验验证平台. *中国科学 E 辑: 信息科学*, 2008, 38(10): 1783-1794)
- [9] P. Szegedi, Deliverable JRA2.1: Architectures for virtual infrastructures, new Internet paradigms and business models, Version 1.6, FEDERICA project, European Union 7th Framework, October 2008
- [10] Xie Gao-Gang, Zhang Yu-Jun, Li Zhen-Yu, Sun Yi, Xie Ying-Ke, Li Zhong-Cheng, Liu Yun-Jie. A survey on future internet architecture. *Chinese Journal of Computers*, 2012, 35 (6): 1094-1108 (in Chinese)
(谢高岗, 张玉军, 李振宇, 孙毅, 谢应科, 李忠诚, 刘韵洁. 未来互联网体系结构研究综述, *计算机学报*, 2012, 35(6): 1109-1119)
- [11] Jacobson V. Special plenary invited short course: (CCN) Content-centric networking, Future Internet Summer School, Bremen, Germany, 2009
- [12] Zhang L, Estrin D et al. Named data networking (NDN) project. PARC Technical Report NDN-0001, 2010
- [13] I. Seskar, K. Nagaraja, S. Nelson, and D. Raychaudhuri, MobilityFirst future internet architecture project. In *Proceedings of AINTEC*. 2011, 1-3
- [14] P. Müller and B. Reuther, Future Internet Architecture - A Service Oriented Approach. In *Proceedings of it - Information Technology*. 2008, 383-389
- [15] Zhang Hong-Ke, Dong Ping, Yang Dong. Theory and key technologies of new generation Internet. *ZTE Communications*, 2008, 14(1): 17-20
(张宏科, 董平, 杨冬. 新互联网体系理论及关键技术. *中兴通讯技术*, 2008, 14(1): 17-20)
- [16] Zhao Liang, Wang Binqiang, Zhang Peng. Research on Architecture of Flexible Reconfiguration Network. *Telecommunications Science*, 2012(2):133-137
(赵靓, 汪斌强, 张鹏. 可重构柔性网络体系研究. *电信科学*, 2012(2):133-137)
- [17] M. Carbone and L. Rizzo, Dummynet revisited. In *Proceedings of Computer Communication Review*. 2010, 12-20
- [18] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner. OpenFlow: Enabling innovation in campus networks. *SIGCOMM Comput. Commun. Rev.*, 38(2):69-74, 2008
- [19] Wang Li-jun, Liu Yong-qiang, Zhang Jian, Research on Future Internet Experimental Technology Based on OpenFlow, *Telecommunications Network Technology*, 2011 (06), 1 - 4
(王丽君, 刘永强, 张健. 基于 OpenFlow 的未来互联网试验技术研究, *电信网技术*, 2011 年 6 月第 6 期: 1-4)
- [20] W. Sebastian, T. Christos, F. Shane, and M. Thomas, Resource Description in Large Scale Heterogeneous Resource Federations, *TRIDENTCOM2011*, Shanghai, China, April 2011
- [21] G. Jourjon, T. Rakotoarivelo, and M. Ott. From Learning to Researching, Ease the Shift through Testbeds. In *Proc. of TridentCom 2010*, volume 46 of LNICST, pages 496-505, Berlin Heidelberg, May 2010
- [22] A. Quereilhac, M. Lacage, C. Freire, T. Turlletti, W. Dabbous, NEPI: An Integration Framework for Network Experimentation, In *Softcom 2011*, September 2011
- [23] L. Peterson, S. Sevinc, S. Baker, et al., PlanetLab Implementation of the Slice-Based Facility Architecture, Draft Version 0.05, June 23, 2009
- [24] T. Rakotoarivelo, M. Ott, G. Jourjon, and I. Seskar, OMF: a control and management framework for networking testbeds. In *Proceedings*

- of Operating Systems Review. 2009, 54-59
- [25] R. Sherwood, G. Gibb, K.K. Yap, G. Appenzeller, M. Casado, N. McKeown, G. Parulkar, FlowVisor: A Network Virtualization Layer, OpenFlow Technical Report, October 2009
- [26] K. Yap, T. Huang, B. Dodson, M.S. Lam, and N. McKeown. Towards software-friendly networks. In Proceedings of ApSys. 2010, 49-54
- [27] I. Baldine, Y. Xin, A. Mandal, P. Ruth, C. Heermann, and J. Chase, ExoGENI: A Multi-domain Infrastructure-as-a-Service Testbed. In Proceedings of TRIDENTCOM. 2012, 97-113.



LIANG Jun-Xue, born in 1982, Ph.D. candidate. liangjx@bupnet.edu.cn. His main research interests include cloud computing and network virtualization.

LIN Zhao-Wen, born in 1979, Ph.D., Lecturer. linzw@bupnet.edu.cn. His main research interests include Future Internet key technology and network security.

MA Yan, born in 1955, Professor. mayan@bupt.edu.cn. His main research interests include Future Internet, IPv6 technology and network security.

Background

The current Internet, which was designed over 40 years ago, is facing unprecedented challenges in many aspects, such as mobility, security, QoS/QoE, and energy efficiency. For these problems, researchers have proposed many new network architecture, concepts, services, protocols, and algorithms to build the new Future Internet, which is different from the existing Internet. However, the fast growth and diversification of the Internet made it extremely difficult to introduce new technologies and protocols backed up with sound experimental validation at realistic size testing environment. So it is important to build a large-scale, realistic, controlled experiment platform, which can be used to test and evaluate any of such new ideas. This work is supported by the National High

Technology Research and Development Program of China (863 Program) (2011AA010704, 2013AA014702) and China Next Generation Internet Project (CNGI Project) (CNGI-12-02-027). The project aims to implement an advanced Future Internet testbed, which can help the researchers to design, deploy and test their new ideas in a realistic and controlled environment. There are many efforts have been devoted to this area and gain a number of achievements in the world, especially in U.S.A and EU. We expect to summary these efforts and achievements, in order to help the researchers understand the key technologies and grasp the main trends in building the Future Internet experiment platform.