



SDN 体系结构与未来网络体系结构创新环境 *

毕 军^{1,2,3}

(1.清华大学信息网络科学与网络空间研究院 北京 100084;

2.清华大学计算机科学与技术系 北京 100084;

3.清华信息科学与技术国家实验室 北京 100084)

摘 要:在讨论互联网体系结构的基础上,介绍未来网络体系结构的主要国际学术研究进展。介绍了软件定义网(SDN)的发展现状,讨论 SDN 的体系结构,提出一种 SDN 体系结构的抽象模型。最后结合国家“863”计划基金资助项目“未来网络体系结构和创新环境”,介绍一种未来网络体系结构创新环境(FINE)体系结构。

关键词:未来网络体系结构;软件定义网;未来网络体系结构创新环境

doi: 10.3969/j.issn.1000-0801.2013.08.002

SDN Architecture and Future Network Innovation Environment

Bi Jun^{1,2,3}

(1. Institute for Network Sciences and Cyberspace, Tsinghua University, Beijing 100084, China;

2. Department of Computer Science, Tsinghua University, Beijing 100084, China;

3. Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing 100084, China)

Abstract: Internet architecture was discussed firstly, then the major international research progress of future network architectures was introduced. And the recent development of software defined networking (SDN) was introduced, an abstract model of SDN architecture was proposed. Finally, a FINE architecture for future network innovation environment project was introduced, which was funded by national high-tech R&D program (“863” program).

Key words: future network architecture, software defined networking, future network innovation environment

1 引言

互联网经过了 40 多年的发展历程,从 1969 年互联网的雏形 ARPANet,发展到 20 世纪 80 年代的学术网络 NSFNet,再到 90 年代开始大规模商业化应用。今天的互联网已经取得了巨大成功,已经发展为世界上覆盖范围最广、规模最大、信息资源极其丰富的社会和经济高度依赖的全球信息基础设施。然而随着互联网的快速发展,当前

互联网面临着许多重大技术挑战,如地址空间濒临枯竭、网络流量重复拥塞、服务质量无法有效保证、网络安全难以根本解决、网络管理手段匮乏等问题。面对这样的问题,首先应该承认互联网的成功是第一位的,应该看到目前遇到的问题是一种“successful disaster”,即飞速发展、高度成功后带来的问题。其背后的技术实质是由于在互联网设计早期,其主要目标不是承载全社会信息基础设施这样的重任,因此随着互联网的大规模使用出现了大量需要解决的

* 国家高技术研究发展计划(“863”计划)基金资助项目(No.2013AA010605),国家自然科学基金资助项目(No.61073172)



问题。

随着互联网使用目标、覆盖范围、应用类型的不断扩大,互联网也通过不断增加新的协议以及对网络体系结构进行各式修补来适应网络的高速发展。这样一种不断自我完善的过程,以国际互联网标准组织 IETF 为主导,通过不断制定新的国际标准来不断增强互联网的功能。比较著名的是 IPv6 的提出。IPv6 虽然对 IPv4 主要的改进仅在地址空间等几个方面,但基于 IPv6 的协议创新仍然在 IETF 不断进行,例如在增强协议安全方面的 SAVI WG(源地址验证提高工作组)^[1]相关标准,解决 IPv4 和 IPv6 过渡的 Softwire WG(软线工作组)^[2]相关标准等。同时,国际标准研究组织 IRTF 也进行面向更长远一些的标准研究,例如最近提出的 IPv6 地址空间同时提供 64 bit 定位符(用于路由和定位)和 64 bit 标识符(用于标识移动的实体)以解决移动性等问题的 ILNP(ID locator network protocol)^[3]等。

但是,这些方法在解决互联网快速发展过程中遇到的各种问题的同时,也加剧了互联网体系结构的复杂性及管理难度。产生这些问题,主要是由当前互联网自身体系结构所决定的,特别是 IP 层作为一种网络体系结构的“窄腰(narrow waist)”,其所承载的重任导致对其进行重大修改是十分艰巨的。所以,设计新型网络体系结构以解决当前网络所存在的问题已经成为学术界、产业界和运营商的迫切需要。于是人们提出一种用 Clean Slate 的方式来研究和设计未来网络体系结构的思路。笔者认为未来网络体系结构的研究与当前互联网的演进不是对立的,所谓 Clean Slate 应该是一种鼓励科研人员大胆探索和设计未来互联网的思维方式,不是说当前互联网应该被推倒重来。新的网络体系结构如何将来能得到实际部署,也必然是从当前互联网逐步过渡而来。另外,未来网络的大规模实际应用仍然被认为需要一个较长的时期,在这个期间,对当前互联网的改进仍然是有重要现实意义的。

在尝试改变互联网的体系结构的同时,软件定义的网络(software defined networking,SDN)被作为一种新型的网络体系结构提出。OpenFlow^[4]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。OpenFlow 实现了将设备控制平面与数据平面相分离,让用户对设备的控制和修改变得更容易。SDN 目前已成为当前全球网络领域最热门的研究方向,在权威机构 IT 领域预测未来 5 年十大关键趋势和技术影响中排名第二^[5]。如果仅看 OpenFlow 技术,似乎与当前网络体系结构没有非常本质

的区别,那么科学的 SDN 体系结构应该是什么呢?

本文首先讨论当前互联网的体系结构特点,然后介绍目前国际学术界对未来网络体系结构的主要研究。接着,在介绍 SDN 的基础上,提出一种 SDN 体系结构模型。最后介绍由清华大学牵头承担的基于 SDN 思想的未来网络体系结构创新环境的研究的进展。

2 互联网体系结构和未来互联网体系结构的学术研究

2.1 互联网体系结构

互联网是人设计的,人当然可以不断改进和优化它。革新互联网的体系结构并不是一件可怕的事情。实际上,互联网的体系结构也不是一成不变的。被称为“互联网之父”的 Cerf V G 博士由于在提出互联网中所做的贡献,于 2004 年获得了被称为计算机科学界诺贝尔奖之称的图灵奖。在 Cerf V G 和 Kahn R E 合作发表于 1974 年的提出互联网的著名论文中^[6],并没有将点到点的路由和转发与终端到终端可靠机制分开。因此互联网早期的核心协议是一种网络控制协议(network control protocol,NCP),即 RFC675^[7]。在实践中人们逐渐总结出“端到端属性”(end-to-end argument)(经总结发表在 1984 年的著名文献^[8]中),即凡是能在高层中实现的功能,不要在低层中实现,因此凡是在端系统(主机)中完成的功能,不应该在网络中完成,这导致了 IP 和 TCP 的“分家”。在 1981 年制定的 RFC791^[9](即 IPv4)和 RFC793^[10]中,将端到端的可靠性机制由 TCP 层来承担,而 TCP 层是在端系统中为应用服务的,路由器不需要承担这样的机制。1983 年 1 月 1 日这天,互联网的前身 ARPANet 用 TCP/IP 取代了旧 NCP,从而成为今天的互联网的基石,这就是今天互联网 IPv4 的由来。

网络体系结构的设计必须满足需求,有什么样的需求就导致设计出什么样的体系结构。在被誉为互联网体系结构设计师的美国麻省理工学院,Clark D 博士总结了互联网体系结构的设计哲学。在 1988 年的著名文献^[11]中介绍了互联网体系结构的最主要目标是把当时各类其他网络连接起来,其他次要目标包括生存性、指出多种服务、适应不同网络、允许分布式管理、效能价格合理、主机易于接入、资源计费等。任何设计都不是万能的,不可能满足所有的目标,互联网体系结构满足了将其他网络给互联起来的主要目标,这也从它的名字“Inter-Net”可以看出,它的本质

是一种“网间网”。至于其他目标,Clark D也坦承不是都满足得很好。更重要的是,随着互联网的日益应用,又涌现了新的当时应用环境和技术条件没有认识到的需求,例如安全性、移动性、可扩展性等。

这种“连接各种不同网络”作为主要需求,就导致了目前这种互联网体系结构设计结果。当前互联网的体系结构一般用图1所示的hourglass(计时器)模型来表示。其中IP层处于最核心的地位,被称为“窄腰”。基于IP over everything的方式,IP层可以在各种底层协议之上构建。另一方面,everything over IP,各种应用可以在IP以及两个重要的传输层TCP和UDP之上构建。笔者认为,互联网体系结构的成功因素就在这里,就是对底层,可以容许和兼容各种通信技术的发展;对上层,可以很方便地促进应用层协议的创新。特别是网络应用开发在主机上进行,而20世纪八九十年代是PC机的时代,人们可以很方便地在PC机上进行网络应用的创新,有了好的应用协议和应用系统,不需要网络设备进行升级,就可以在端系统上部署,而且只需要部署部分主机就可以马上开始应用,随着市场的选择来促使更多的用户来安装新的网络应用协议和应用系统,这催生了90年代以来的众多互联网公司创新。同时,这种IP“窄腰”,由于它太重要了,每个网络设备和主机都要安装,因此IP的改进就要求设备和主机的升级。这种IP层的“封闭性”和对网络应用的“开放性”导致了互联网应用的繁荣和IP层缓慢演进的鲜明对比,虽然每年对IP核心层的研究成果非常多,但能对IP核心层的改变发挥实际作用的技术十分少,只有极少数专家才能设计出对IP核心层进行合理改进的RFC标准。

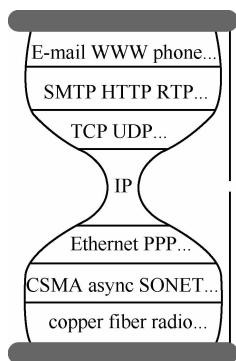


图1 当前互联网体系结构的“计时器”模型

今天,再一次站在了网络体系结构设计的十字路口上。如果有像当年Clark D收到DARP给他设计互联网体系结构时的需求这样的机会,要提出什么样的设计目标

呢?今天,连接不同的网络虽然仍然重要,但可能其重要性不能说是首要目标了。今天已经在互联网的实践中面临了许多当年想不到的重要需求,例如安全性、可扩展性等。另外,已经从PC机时代发展到云计算时代、移动计算时代。为了适应时代的需求,更好的管理、控制和定制能力,服务的快速引入(类似苹果安卓这类加载模式),移动性,新的网络经济模式等需求变得更加重要。最重要的是,已经看到了需求是不断变化的,对未来需求的预测总是有限的,则很需要应对改变的能力,即可演进能力和平滑革新的能力。笔者个人观点认为如果提出新的网络体系结构必须满足最重要的需求,那么通过开放性来提供可定制和可演进能力的需求是最重要的,主要理由是除了适合于目前云计算时代的要求,也为未来需求提供了平滑演进的能力。SDN正是为人们提供了实现这种能力的一个方向,受到了人们的高度重视。

2.2 未来互联网体系结构的国际研究进展

国际学术界试图针对长远的目标来研究互联网的未来发展及其新技术。这个领域目前一般被称为“未来互联网”,包括研究、探索和试验等工作,是一个日益受到研究界和工业界关注的十分活跃的新领域。为此,北美、欧洲、日韩等许多国家和地区都十分重视未来网络体系结构的研究和未来网络创新环境的建设,启动了FIND、GENI、FIA、FIRE、4WARD、AKARI等科研项目。我国也在“973”、“863”、国家自然科学基金等项目中对未来网络研究进行了资助。

由于篇幅关系,本节主要介绍美国国家自然科学基金会(NSF)资助的未来互联网体系结构的研究情况。其归纳为如下几个阶段:

- NewArch项目(2000–2003年);
- FIND计划(2005–2009年);
- FIA计划(2010–2013年);
- 新一期FIA-NP计划(2013–2015年)。

NewArch^[12]项目是美国国防部高等研究计划局(DARP)资助的Future-Generation Internet Architecture,由MIT的Clark D负责,参加单位包括USC/ISI、MIT LCS、ICSI等美国大学和研究所。NewArch项目的主要成果包括:FARA(forwarding directive, association, and rendezvous architecture,转发指令—关联—约会体系结构)模型,可以认为是尝试改变当前互联网IP语义中定位符和标识符的“重载”(overload)的,把基于IP地址的路由变为基于FD的路由,



把基于 IP 地址+TCP 端口做接收方标识变为基于 AID (association ID)做接收方标识,通过约会服务把标识符和定位符关联起来;RBA(role based architecture,基于角色的可组装体系结构)尝试把当前互联网体系结构的 building block 从基于协议变为基于角色,通过“角色”来组建网络通信;NIRA (new internet routing architecture)——用户控制的网络层路由,尝试改变基于 ISP 选择的域间路由为由用户选择的域间路由结构,增强用户的控制力;XCP(eXplicit control protocol,明确控制协议)尝试把互联网的端到端拥塞控制变为路由器参与的拥塞控制以提供新的 QoS 框架的。

NerArch 的研究成果虽然没有产生实用价值,但它促进了一批学者产生对未来互联网体系结构的思考和浓厚兴趣,因此美国 NSF 于 2005 年开始资助了 FIND 计划^[13]。FIND 计划分两期共资助了 50 余个未来互联网研究项目,例如通过核心与边缘分离来提供域间路由扩展性的 eFIT 项目,引起了 IRTF 的 RRG(路由研究组)开始进行路由可扩展性研究,收到了 10 余项技术;接着 IETF 成立了 LISP 工作组,提出了 ID 和 Locator 分离的标准协议。虽然 eFIT 项目本身最终没有产生重要影响,但对促进网络界关注通过核心与边缘分离技术解决域间路由扩展性问题,通过 ID 和 Locator 分离解决移动性问题产生了一定的影响。由于研究未来互联网体系结构的人越来越多,FIND 计划虽然结束了,但在 NeTS 计划中,仍然接受“FIND-like”类的项目,鼓励研究者继续进行未来网络体系结构的小规模探索类项目。

在 FIND 项目的基础上,NSF 认为有必要加大资助力度,因此于 2010 年启动了 FIA(future internet architectures)计划^[14],通过两轮筛选,于 2010 年 8 月遴选出了 4 个准备比较充分的项目,每个项目资助了 700~800 万美元不等。这 4 个项目是目前网络体系结构界关注的焦点之一。UCLA 大学 Lixia Zhang 教授负责的 NDN(name data networking)项目^[15],主要提出一种内容中心网的体系结构,变“IP 为中心”为“内容为中心”,变基于 IP 地址的路由为基于内容名字(name)的路由以及增加了安全性等功能。Rutgers 大学 Dipankar Raychaudhuri 教授负责的 MobilityFirst 项目^[16],主要是针对移动性的需求,提出了全球唯一的标识符,与网络地址进行映射服务;与 NDN 有机构的名字不同,Mobilityfirst 的标识符是一种平坦的 ID,标识符生成时可以自带认证能力。宾夕法尼亚大学 Smith J 教授负责的 NEBULA 项目^[17],主要针对云网络的需求设计其体系结构。

CMU 大学 Steenkiste P 教授负责的 XIA(eXpressive internet architecture)项目^[18],提出构建一种新型的“窄腰”来应用不同的网络需求。另外,麻省大学 Wolf T 教授负责的 ChoiceNet 项目^[19]也于 2012 年从 NeTS 计划里纳入了 FIA 计划,成为第 5 个 FIA 项目,该项目期望对同样的网络需求有不同的实现,然后通过市场选择机制和技术支撑,为用户提供选择后的服务。前面 4 个项目经过 3 年的研究,已经硕果累累,目前已经完成了项目总结,正在争取新一期 FIA 资助。ChoiceNet 由于启动较晚,项目仍然在进行中。

在 FIA 完成的基础上,2013 年 NSF 又针对新一期 FIA 计划(FIA-NP)^[20]进行了项目招标,但是要求申请人必须是 FIA 项目里的成员。2013 年 6 月已经完成了项目申请书的提交,马上就可以看到其评审结果。

3 SDN 的发展现状

在 2010 年 FIA 进行项目申请时,NSF 提出了所有项目申请书需要解决的问题。其中必须解决的问题是安全可信性(security, privacy, reliability, and usability-must be considered as a fundamental design requirement in proposed architectures),即安全、隐私、可靠性、可用性等必须作为所提出新体系结构的设计基础。FIA 项目其他可选的要求包括可扩展性(scalability)、开放性(openness)、无所不在的访问(ubiquitous access)、支持创新(innovation-enabling)、可管理型(manageability)、可演进性(evolveability)、经济活力(economic viability)、开发原型和试验网评价(在 GENI 等试验评价)等。

2010 年斯坦福大学牵头也将 OpenFlow 为基础的软件定义网进行了 FIA 项目申报,但是在第 1 轮就被淘汰了。主要的原因是项目没有很好地对照标书应答,例如在安全可信性以及各项需求方面没有给出很好的应答。2010 年笔者曾就此问题与斯坦福大学的 McKoean N 教授交流过,他认为当时如果申请书中基于 OpenFlow 运行一个具体支持安全性等符合需求的新体系结构,就可以符合要求。这也说明,SDN 应该是与上述 FIA 各体系结构不是同一个层面的技术,SDN 应该是一种更基础的体系结构。

SDN 的定义目前还没有达成共识,工业界的 ONF、国际标准界的 IRTF SDN RG、国际学术界的各个学术会议给予了不同的定义。一般认为,SDN 通过将网络控制与网络转发解耦合构建了开放可编程的网络体系结构,但具体技

技术和SDN体系结构的重点所在仍然在认识上差别较大。如果一个技术没有定义,则仅能成为思想或领域,所以笔者认为SDN与其说是具体的体系结构,不如说是一种开放式网络的思想、目标,是一个研究领域。

虽然业界已经基本达成共识,OpenFlow技术不等同于SDN,OpenFlow是SDN体系结构里的南向接口的实例,OpenFlow的重要性在不同技术流派中也不同,但目前比较具体清晰的技术仍然只是基于OpenFlow的SDN体系结构,并且得到了实际使用。

OpenFlow技术可以追溯到斯坦福大学于2007年SIGCOMM国际学术会议上发表的论文Ethane^[21]。这个研究中,斯坦福大学感到对校园网的网络设备进行开放控制是一件很有价值的工作,于是开展了基于校园网设备支持各种研究的实验,于2008年在ACM CCR期刊上正式发表了OpenFlow^[22],提出了OpenFlow可以在校园网上通过不同的VLAN支持不同的新协议的研究和创新。OpenFlow的特点是由交换机外的控制器上运行的应用程序来控制交换机内的十几个报文头的域组成的流表,以此定义网络设备行为的方式。这篇仅有6页的短文不像一般学术论文那样充满了公式和实验数据图表,但是却被工业界高度重视,并反过来促进了学术界跟进,继续深入研究。OFELIA^[23]于2010年开始采用具有标准化接口的OpenFlow技术搭建统一新一代网络实验平台。GENI和FIRE在2011年提出将采用OpenFlow技术构建下一代网体系结构的创新环境。同年,开放网络基金会(Open Networking Foundation, ONF)成立,致力于软件定义网络的发展和标准化^[24],2013年组成了以厂商为主的、进行开源软件开发的Open Daylight Project^[25]等组织。国际互联网研究的顶级学术会议SIGCOMM于2012年8月成立HotSDN Workshop(2013年继续举办),2013年国际顶级会议也把SDN纳入了征稿范围。与IETF一起活动的国际互联网研究组织IRTF于2012年10月成立了SDN研究组(SDNRG)^[26]。

目前OpenFlow的标准化工作由ONF负责进行,因此ONF也是SDN界目前最重要的组织,2013年对SDN的发展起到了重要推动作用。为了避免受到网络设备厂商的控制,ONF确定以用户为主导的原则,因此其理事会成员包括分处于三大洲的运营商Verizon、德国电信、NTT以及Google、Facebook、Microsoft、Yahoo、高盛等用户界代表。目前ONF的成员已经达到100个,包括运营商、互联网公司、IT厂商、设备厂商等。ONF包括如下8个工作组。

- Extensibility工作组负责OpenFlow标准版本的维护。
- Configuration & Management工作组负责OF-Config配置和管理标准的维护。
- Testing & Interoperability负责测试工作。
- Architecture & Framework工作组负责相关体系结构和术语的文档制定。
- Forwarding Abstractions工作组负责尝试新的转发抽象技术如TTP。
- Optical Transport工作组负责将这些技术应用到光传输领域。
- Migration工作组负责研究如何从现网过渡到新的OpenFlow网络。
- Market Education工作组主要负责与市场合作的宣介工作的文档。

另外,ONF还有若干的讨论组,包括Security、Wireless and Mobile、SDN Future等。为了加强与研究界的合作,ONF最近还从全球各大学和研究机构邀请了13位研究顾问。

OpenFlow的标准版本1.0于2009年12月发布,其最主要特征是:基于12元组match(in_port + L2, L3, L4)进行3类action(forward, modify header, send to controller),非常容易用现有硬件实现,所以目前业界能提供的OpenFlow交换机主要支持OpenFlow 1.0标准。OpenFlow 1.1于2011年2月发布,最主要特征是多流表流水线,值得注意的是OpenFlow 1.1与OpenFlow 1.0不能实现后向兼容,因此从OpenFlow 1.1开始,OpenFlow标准有了两个分支,虽然目前制定了新的OpenFlow标准版本,但仍然要维护OpenFlow 1.0.1版本。OpenFlow 1.2于2011年12月发布,其最主要特征是通过OXM来支持灵活性。OpenFlow 1.3于2012年4月发布,近期又发布了OpenFlow 1.3.1和1.3.2等,对OpenFlow 1.2和OpenFlow 1.1都具有后向兼容性。OpenFlow 1.4标准目前已经在ONF内部审阅,估计2013年8月初获得批准发布。OpenFlow 1.5标准计划2013年底或2014年初批准发布。

在学术研究界,SDN目前学术研究十分活跃,笔者初步统计在SIGCOMM、NSDI、OSDI、CoNEXT等计算机网络的国际顶级学术会议上,已经有20余篇关于SDN的学术论文,这些研究主要集中在控制器的设计、新型数据平面的设计、编程语言和接口、测试和调试、应用(包括数据中心的應用、校园网的应用、网络安全的应用、MiddleBox的



应用)等。特别是2013年国际网络学术界最高水平的学术会议SIGCOMM2013上就有约1/5的论文是与SDN相关的学术研究。从另一个角度看,这也说明SDN技术仍然在发展中,还有很多基础层面需要解决的问题,例如什么是科学的SDN体系结构、如何提供更强的数据平面抽象能力、如何提高SDN的可编程性和接口、SDN的网络可扩展性和性能、SDN的安全性、SDN的管理和测量、SDN的应用案例分析等。

但SDN也不仅停留在学术研究层面,应该说工业界的热情不亚于学术界,谷歌、微软等互联网公司均在SDN领域投入了大量的科研力量,思科、Juniper、爱立信、阿朗、IBM、HP、Dell、NEC、Arista、NTT、Big Switch、Ixia等IT厂商也正在研制SDN控制器和交换机。我国华为技术、中兴通讯、华三、神州数码、锐捷、盛科等企业在此领域的开发和产品化也取得了进展,成立SDN与开放式网络专委会^[27]进行SDN技术的推进。特别是很多互联网公司、IT设备厂商把SDN当作一次重新洗牌的机会,积极投入了力量。目前也已经开始有工业界的成功案例。虽然前面谈到OpenFlow技术还需要完善,其标准仍然在不断发展之中,但已经对SDN的应用起到了推动作用,例如产生了Google B4^[28]和微软SWAN(software-driven WAN)^[29]等一些成功的应用案例。谷歌和微软设计优秀的流量工程算法,通过OpenFlow接口控制开放式网络设备,极大提高了跨数据中心私有WAN的链路利用率。Google互联了跨欧亚三大洲的12个数据中心网,微软互联了跨欧亚三大洲的5个数据中心网。它们均声称将跨数据中心的链路利用率提高到了接近最优:Google声称达到了接近100%,而微软声称从使用MPLS-TE的58.3%提高到99%,还提高了流的公平性

和网络更新的相应速度等。

应该承认,SDN的实际应用仍然存在一些挑战,SDN的实际应用案例还不多。笔者认为这与SDN对用户的技术能力和创新意识要求比较高有一定关系。传统的用户可能愿意等待厂商把一切都做好了,没有意识到未来的网络其实就应该是用户自己来定义和实现(这个转变如同从软件的使用者转变到软件的编程者)。现在很多厂商已经开始技术转型,把设备和网络开放给用户来控制;但用户是否做好了技术转型的心理和技术准备?因此笔者估计,那些有较高技术开发能力和创新意识的公司(如前面说的Google和微软这类创新意识很强的互联网公司以及具有开发能力的各类网络用户,特别是行业专网和大型企业网的用户),将是SDN新技术的最先的使用者。然后随着用户获得收益出现了更多普通用户的需求,会逐渐形成一个新的产业链,即出现提供SDN专项服务(例如配套的软件、工具和接口,应用设计,技术服务等)的新型企业,帮助那些尚不具备自主创新能力的普通用户来应用SDN。当然,出现这个新的产业链需要较长的时期。在这个期间,有SDN需求但不了解如何入手的企业,可以考虑与设备厂商和高校合作,根据用户的实际需求,结合高校的科研力量和厂商的服务,一起开发新的SDN应用。

4 SDN的体系结构及其抽象模型研究

软件定义网的体系结构一般被认为是如图2所示的结构,包括了网络应用层、控制器组成的控制层、交换机组成的基础设施层3个重要的组成部分。该图的来源是ONF,因此笔者称其为基于OpenFlow的SDN体系结构。如果研究Open Daylight项目或IRTF SDNRG,将发现其SDN

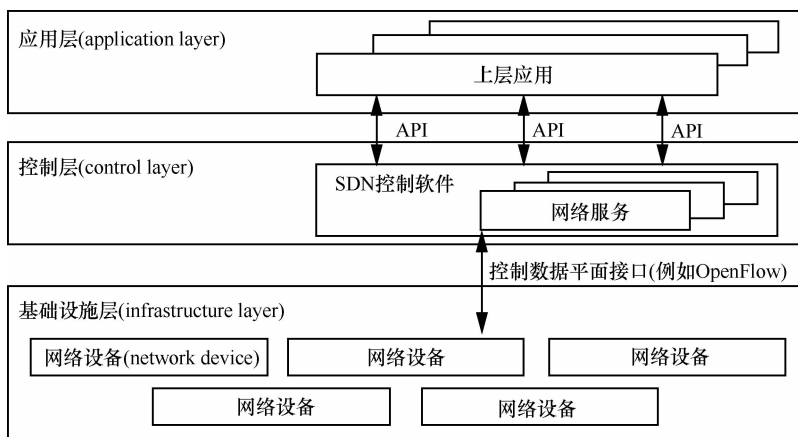


图2 基于OpenFlow的SDN体系结构

体系结构与之有不同之处。那么到底什么是 SDN, 什么是 SDN 的体系结构呢?

虽然前面谈到基于 OpenFlow 的 SDN 得到了学术界和工业界的重视, 也取得了成功, 但这个体系结构没有回答一个重要的问题, 即如果第 2 节中分析的开放性是新型网络体系结构的最重要需求, SDN 的结构显然支持了开放性 (例如把设备打开给控制器控制, 最终由应用来控制), 但为什么图 2 的结构就是一种科学的网络体系结构设计。下面谈一下笔者的个人认识。

笔者认为开放性的实质在于有一种对网络功能和其他网络体系结构的表述能力, OpenFlow 提供了基于 Layer 2 (MAC 层各域)、Layer 2.5 (MPLS)、Layer 3 (IP 各域)、Layer 4 (传输层各域) 功能如何使用的灵活定义, 从而支持开放性和演进性的需求, 这是其获得认可的基础。但是基于 OpenFlow 的 SDN 显然只能对上述协议域进行控制, 仍然有一定局限性。SDN 是一种基础层次的技术, 基于 SDN 可以进行各种新型网络体系结构的研究和设计, 基于 OpenFlow 的 SDN 离这样的目标仍然有一定差距, 还不能反映 SDN 的实质。

基于上述分析, 笔者认为: SDN 是提供对各种体系结构和协议进行描述的能力, 是一种“元体系结构 (meta architecture)”。如同元数据 (meta data) 是对数据进行描述的数据, 元语言是对语言进行描述的语言, 则元体系结构是一种对体系结构进行描述的体系结构。这种表述能力支撑了 SDN 可以对各种体系结构描述, 并可以基于 SDN 的架构支持演进。也正因为 SDN 是一种元体系结构, 其描述能力是可以有不同定义的, 因此出现在 SDN 思想下的各种不同技术途径是正常的, 根据被描述对象的不同就允许有不同的 SDN 技术和结构。OpenFlow 仅是 SDN 描述数据平面的一种技术, 不能等于全部, 人们可以大胆地设计其他实现这种描述的技术路线, 都应该被认为是 SDN 的体系结构。

为了使各种不同的 SDN 体系结构设计能有一定科学规律可循, 最近提出了一种 SDN 系统结构抽象模型, 如图 3 所示。该模型的出发点是认为 SDN 所提供的开放性的实质是: 不应无限制地增加网络的复杂度, 需要对网络进行抽象以屏蔽底层复杂度, 为上层提供简单的、高效的配置与管理。

在图 3 中, 低层网络为高层网络提供网络状态的抽象, 中层根据这个网络状态抽象以及本层的状态, 在上层的控制命令控制下, 生成对下层的控制命令以及本层的新

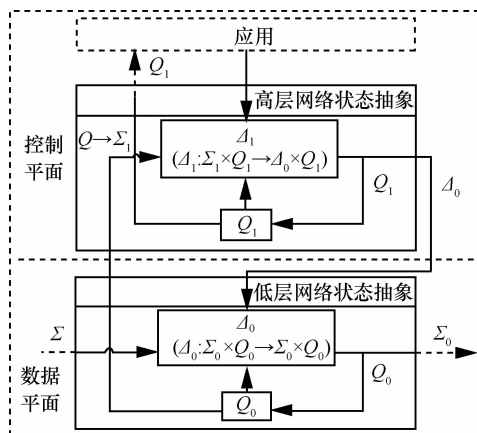


图 3 SDN 体系结构的抽象模型

状态。这种上下层之间的网络状态抽象接口, 上下层之间的控制命令接口, 就是 SDN 的可编程接口或“软件定义”的接口。理论上讲, 这样的结构可以是不断递归的, 从而实现通过递归的方式“降解”网络复杂度, 达到用户可以通过高级接口容易地设计和实现网络功能的目的。因此基于该模型, 可望建立一种新的基于复杂网络系统的抽象来研究、设计、管理、运行网络的网络科学, 使今天的靠设计哲学、设计经验来设计和运行网络, 变成靠一整套计算机科学体系和工具来支撑。所以 SDN 模型与 IP 网络是不矛盾的, 其实质是一种对网络的抽象和描述能力, 可促进网络的研究和设计。基于 SDN 既可以进行新型网络体系结构的研究和设计, 也可以进行 IP 网络新协议的研究和设计。

基于该抽象模型可以实例化出基于 OpenFlow 的 SDN 网络体系结构。令 $\Sigma = \{\Sigma_0, \Sigma_1\}$ 为输入的有限集合, 其中 Σ_0 为数据平面的输入, Σ_1 为控制平面的输入。 $Q = \{Q_0, Q_1\}$ 是状态的非空有穷集合, 其中 Q_0 为数据平面的状态, Q_1 为控制平面的状态。 $\Delta = \{\Delta_0, \Delta_1\}$ 是状态转移函数, 其中 Δ_0 为数据平面的转移函数, Δ_1 为控制平面的转移函数。应用通过高层网络状态抽象将其需求通过 Δ_1 传递给控制平面。控制平面根据数据平面的状态 Q_0 映射为控制平面的输入 Σ_1 , 通过控制平面转移函数 $\Delta_1: \Sigma_1 \times Q_1 \rightarrow \Delta_0 \times Q_1$ 生成下一层的转移函数 Δ_0 并将控制平面新的状态 Q_1 返回给应用。数据平面通过控制平面下发的转移函数 Δ_0 , 将数据平面的输入 Σ_0 和数据平面状态 Q_0 利用转移函数 $\Delta_0: \Sigma_0 \times Q_0 \rightarrow \Sigma_0 \times Q_0$ 转换为数据平面的输出 Σ_0 , 并更新报文处理后的数据平面状态 Q_0 。其中这种转移函数就是网络应用对 OpenFlow 流表的各类灵活控制, 也就是前面提到的 SDN 的价值在于所提供的“描述能力”。

也可以基于该抽象模型实例化出其他 SDN 体系结构,



例如不一定是基于 OpenFlow 的 SDN 这样 3 层结构,也可以递归出很多的层次。如同今天一些高级的网络编程语言已经在很高的层面编程,SDN 也可以有很高级的编程接口,供那些不希望解除网络底层的用户来描述自己所需的网络体系结构或网络协议功能。通过这种的模式,未来可能改变今天“网络设计靠高人、网络管理高老手”的模式,使更多的年轻人参与网络体系结合和协议的设计,得到实践,达到今天网络应用层创新的繁荣程度,促进网络核心技术的创新。

因此,基于这种抽象模型,也可以实例化出一种未来网络体系结构创新环境(FINE)的 4 层体系结构。

5 未来网络体系结构创新环境网络体系结构

清华大学 2010 年与斯坦福大学签署了未来网络研究的合作协议后,即开展 SDN 相关的研究,先后与华为技术、中兴通讯、神州数码、惠普等企业进行了合作研究,近几年研究的领域包括体系结构、OpenFlow 扩展、新型网络设备、控制器、测试、应用等方面。2012 年清华大学牵头的国家“863”计划项目“未来网络体系结构和创新环境”获得了批准。该项目由清华大学牵头负责,清华大学、中国科学院计算所、北京邮电大学、东南大学、北京大学等分别负责各课题,笔者担任项目首席专家。项目参加单位还包括解放军信息工程大学、工业和信息化部电信研究院、国防科学技术大学、上海未来宽带技术及应用工程研究中心有限公司、总参 61 所、中兴通讯、华三、华为技术、锐捷、神州数码、中科院网络中心、西安交通大学、北京市天元网络科技股份有限公司、中国科学院声学所、上海数字电视国家工程中心、北大方正等。项目的出发点在于:新型网络体系结构和新型网络协议在设计方法、编址方式、转发机制、控制

模式等方面存在巨大差异,而当前网络设备和试验环境的封闭性又严重制约着网络的技术创新和体系演进,因此,设计和实现促进未来网络体系结构创新的环境,对支撑未来网络的技术创新和体系演进具有重要的意义。项目的目标是:采用创新理念和技术路线,研究未来网络创新体系结构,突破关键技术,研制新型网络设备和软件系统,建设未来网络体系结构的创新试验环境,研究内容中心网等各种新型网络体系结构和新协议,依托 FINE 进行实验验证。

前面提到 SDN 的核心价值在于具有对其他体系结构进行描述能力的“元体系结构”,因此选择了 SDN 思想来作为未来网络体系结构创新环境的设计思想。前面也提到项目提出了未来网络体系结构创新环境(FINE)体系结构,如图 4 所示。FINE 体系结构包括 4 个层次,即数据平面抽象和开放设备层、网域操作系统层、虚拟化平台层、新体系结构和新协议层。数据平面抽象和开放设备层提供设备本地视图的编程接口供上层控制;网域操作系统层对上层提供全局的物理视图编程接口,方便上层(新体系结构或新协议可以直接编程或通过虚拟化平台层来控制);虚拟化平台层对上层提供特定的逻辑视图编程接口,为应用层提供虚拟化资源;还有一个比较重要的部分是域间协商通信机制,跨域不应该是一种控制机制,而应该是一种不同管理域的域间协商机制,通过网域操作系统的“东西向”接口,实现对邻居网络的通信链路、路由、资源和服务质量的协商。

基于 FINE 体系结构,项目通过研究和设计各种不同的数据平面转发抽象技术及其开放式网络设备(包括 OpenFlow 的扩展技术以及其他数据平面抽象技术),新型网域操作系统(与 OpenFlow 控制器这种“通道方式”相比,主要是提供了应用与设备的隔离以及资源的管理)和虚拟化平台(主要是虚拟资源与物理资源的映射),支撑各种新

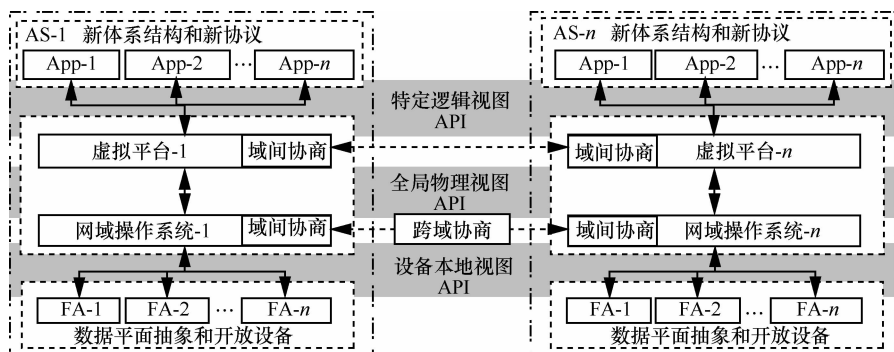


图 4 FINE 体系结构

型网络体系结构(例如新型内容中心网)和IPv6新协议(例如IPv6真实源地址验证)的研究试验。项目在研究和开发软硬件系统的基础上,依托教育网和各参加单位的网络建设FINE的试验平台。这方面的具体技术内容由于篇幅原因,将在其他文章中介绍。

6 结束语

未来网络体系结构已经成为人们关注的热点之一,近年来SDN更获得学术界和工业界的高度重视。本文从网络体系结构的设计需求角度讨论了互联网体系结构,介绍了未来网络体系结构的国际主要学术研究进展、SDN的发展现状,讨论了SDN的体系结构,提出了SDN体系结构的抽象模型,结合笔者正在负责承担的国家“863”计划“未来网络体系结构和创新环境”项目,介绍了一种未来网络体系结构创新环境(FINE)体系结构。

致谢:感谢清华大学信息网络科学与网络空间研究院网络体系结构和IPv6研究室的同事姚广博士、胡虹雨博士、王旻旻博士等,以及我所指导的博士生冯涛、付永红、朱树勇、高凯、林萍萍、刘冰洋、吴承辉等,他们与我进行的多次学术讨论对我此文的撰写起到了重要作用。也感谢国家“863”计划项目“未来网络体系结构和创新环境”的各参加单位的老师和同学们。

参考文献

- 1 IETF SAVI Working Group. <http://datatracker.ietf.org/wg/savi/>, 2013
- 2 IETF Software Working Group. <http://datatracker.ietf.org/wg/software/>, 2013
- 3 Atkinson R J, Bhatti S N. RFC 6740: Identifier-Locator Network Protocol (ILNP) Architectural Description, November 2012
- 4 OpenFlow. <http://www.openflow.org/>
- 5 http://www.cnw.com.cn/news-report/hm2012/20121024_256567.shtml
- 6 Cerf V G, Kahn R E. A protocol for packet network intercommunication. IEEE Trans on Communications, 1974, comm-22(5)
- 7 Cerf V, Dalal Y, Sunshine C. RFC675: Specification of Internet Transmission Control Program, 1974
- 8 Saltzer J H, Reed D P, Clark D D. End-to-end arguments in system design. ACM Transactions on Computer Sysmtes, 1984, 2 (4):227~288
- 9 RFC791: Internet Protocol-DARPA Internet Program Protocol Specification, September 1981
- 10 RFC793: Transmission Control Protocol-DARPA Internet Program Protocol Specification, September 1981
- 11 Clark D. The design philosophy of the DARPA internet protocols, ACM SIGCOMM Computer Communication Review, 1988, 18(4):106~114
- 12 Clark D, Sollins K, Wroclawski J, *et al.* New Arch: Future Generation Internet Architecture. Technical Project Report, 2003
- 13 FIND Project. <http://www.nets-find.net/>, 2013
- 14 FIA Project. <http://www.nets-fia.net/>, 2013
- 15 NDN Project. <http://www.named-data.net/>, 2013
- 16 MobilityFirst Project. <http://mobilityfirst.winlab.rutgers.edu/>, 2013
- 17 NEBULA Project, <http://nebula-fia.org/>, 2013
- 18 XIA Project, <http://www.cs.cmu.edu/~xia/>, 2013
- 19 ChoiceNet Project. <https://code.renci.org/gf/project/choicenet/>, 2013
- 20 FIA-NP Project. <http://www.nsf.gov/pubs/2013/nsf13538/nsf13538.htm>, 2013
- 21 Casado M, Freedman M J, Pettit J, *et al.* Ethane: taking control of the enterprise. ACM SIGCOMM Computer Communication Review, 2007, 37(4):1~12
- 22 McKeown N, Anderson T, Balakrishnan H, *et al.* OpenFlow: Enabling Innovation in Campus Networks. ACM SIGCOMM Computer Communication Review, 2008, 38(2):69~74
- 23 OFELIA Project. <http://www.fp7-ofelia.eu/>, 2013
- 24 Open Networking Foundation. <https://www.opennetworking.org/>, 2013
- 25 Open Daylight Project. <http://www.opendaylight.org/>, 2013
- 26 IRTF Software Defined Networking Research Group. <http://tools.ietf.org/group/irtf/trac/wiki/sdnrg#>, 2013
- 27 中国SDN与开放式网络专委会. <http://www.chinasdn.org/>, 2013
- 28 Jain S, Kumar A, Mandal S, *et al.* B4: experience with a globally-deployed software defined WAN. Proceedings of ACM SIGCOMM, Hong Kong, China, 2013
- 29 Chi-Yao Hong (UIUC), Kandula S, Mahajan R, *et al.* Achieving High Utilization with Software-Driven WAN. Proceedings of ACM SIGCOMM, Hong Kong, China, 2013

[作者简介]



毕军,男,于清华大学计算机系获博士学位。曾赴美国留学,任贝尔实验室博士后、研究员。2003年由清华大学引进回国工作,现任清华大学信息网络科学与网络空间研究院(原清华大学网络研究中心)网络体系结构和IPv6研究室主任、教授、博士生导师,国家“863”计划“未来网络体系结构和创新环境”项目首席专家。目前主要从事新型互联网协议和新型网络体系结构领域的研究工作。

(收稿日期:2013-08-01)