

OpenFlow – 打造弹性化的可控互联网

Tuesday, November 2 2010, 11:00 AM

注：2010/11/07发表于[弯曲评论](#)。

2012的故事

2012年的某天，你跟往常一样起床，打开电脑，却发现无法登录到邮箱、无法连接到公司的VPN网络、无法订购任何东西……，你会发现一切都简直跟世界末日一样，离开了网络，现代社会根本无法正常运行。这并不是可笑的无根据的幻想，如果世界末日真的来临，摧毁互联网无疑是最直接有效的办法，而现代互联网并没有我们想象的那样鲁棒。

从上个世纪70年代初，互联网在短短不到40年时间里已经发展成为这个地球上不可或缺的基础设施。然而由于一开始的设计并没有考虑到后来互联网的规模会如此庞大、承载的应用会如此复杂、地位会变得如此重要，现代的互联网在过重的压力下已经凸显出太多亟待解决的问题。互联网太危险，恶意攻击、病毒、木马每年造成上千亿刀的损失；互联网太脆弱，无标度（Scale-free）的特性让整个网络可以在精心设计的少数攻击下即告崩溃；互联网太随意，p2p等应用的出现一度造成各大ISP网络堵塞，严重影响传统正常的访问；互联网太迟钝，现代臃肿的路由机制不能支持快速的更新，即便发现问题也无法快速反应；互联网太局促，IPv4的分配地址已经捉襟见肘……

这一切的问题都隐隐的指向了互联网这个庞然大物最关键的软肋——可控性。缺乏有效的控制措施让互联网这个为服务人类而设计的机器，正在逐渐演变成一头臃肿而暴躁的凶兽，挣扎着要摆脱人类所施加的脆弱枷锁。

下一代互联网和GENI

为了解决当前互联网的问题，不少国家都纷纷提出了下一代互联网计划，代表性计划有美国的FIND（Future Internet Network Design，未来互联网网络设计）和GENI（Global Environment for Network Innovations，全球网络创新环境），欧洲的FIRE（Future Internet Research and Experimentation，未来互联网研究和实验），中国的CNGI-CERNET（China Next Generation Internet）。所有这些计划参与者大都是各个国家产、学、研顶尖的机构。

这三大计划中，CNGI-CERNET主要是研究在IPv6体系下的新一代网络；而NSF支持的FIND计划计划在不受当前互联网的制约下提出未来互联网的需求，从2006年到2014年分三个阶段主要致力于五个问题：是否继续采用分组交换、是否要改变端对端原理、是否要分开路由和包转发、拥塞控制跟资源管理、身份认证和路由问题。FIND计划最主要的成果之一就是GENI——一套网络研究的基础平台，同时FIRE计划跟GENI项目合作也非常密切。GENI计划的两大任务是为最前沿的网络科学工程领域革命性研究开路；刺激和促进重大社会经济影响的奠基性创新的出现；围绕这两大任务，GENI致力于打造下一代互联网的虚拟实验室，为研究者提供验证创新的架构、协议的灵活、可扩展、可配置的实验平台，并促进学术界和业界的相互合作。长期以来，缺乏合适的实验平台让各界的专家学者们伤透了脑筋，PlanetLab的种种局限已经不能满足广大researcher越来越令人fz的需求了。

毫无疑问，GENI的目标将让每个网络研究者为之着迷和激动，一套完全可控、可定制、大规模的网络试验床，对学术界将意味着大批的顶级paper，对业界意味着大量的新标准、新协议。

OpenFlow的前世今生

GENI的好处虽多，但要部署这个平台无疑是一件太过昂贵的事情，于是一个自然的事情就是在目前现有的网络下，能否省时省力的干好这个事情？

很自然的想法，如果我能控制整个Internet就好了，而网络中最关键的节点就是交换设备。控制了交换设备就如同控制了城市交通系统中的红绿灯一样，所有的流量就可以乖乖听话，为我所用。然而现代交换设备被几家巨头垄断，开放的接口十分有限，能做的事情也十分有限。如果能有一套开放接口、支持控制的交换标准该多好？OpenFlow应运而生。

最初的想法其实十分简单，无论是交换机还是路由器，最核心的信息都存放在所谓的flow table里面，用来实现各种各样的功能，诸如转发、统计、过滤等。flow table结构的设计很大程度上体现了各个厂家的独特风格。OpenFlow就是试图提出这样一个通用的flow table设计，能够满足大家不同的需求，同时这个flow table支持远程的访问和控制，从而达到控制流量的目的。

具体来说，OpenFlow的flow table中每一个entry支持3个部分：规则，操作跟状态。规则无非是用来定义flow，OpenFlow里flow定义十分宽泛，支持10个域（除了传统的7元组之外增加了交换端口、以太网类型、Vlan ID）；操作就是转发、丢弃等行为，状态部分则是主要用来做流量的统计。在此基础上最关键的特性就是支持远端的控制，试想，如果我要改变entry就必须跑到交换机前重新编程写入得多麻烦，而且如果我想获知网络的实时状态咋办，有了统一的控制机制，我们的网络才变得真正智能可控起来。OpenFlow的控制机制也十分灵活，感兴趣的同仁可以参考NOX。

好了，有了这个标准，只要大家以后生产的交换设备都支持，那么学术界以后能做的事情就太多了，以前YY无数次的梦想终于开始变成了现实。比如我们可以在正常运行的网络中自己在定义一些特殊的规则，让符合规则的流量按照我们的需求走任意的路径，就仿佛将一张物理网络切成了若干不同的虚拟网络一样，同时运行而又各不干扰，我们可以轻而易举的测试各种新的协议；以前要做什么处理，需要考虑到具体的拓扑结构，考虑到box的先后顺序，现在好了，通过定义不同的flow entry就可以任意改变流量的运行策略，这也很好的为解决移动性问题提供了便利（一个著名的demo是笔记本在不同交换机之间切换，虚拟机在两地之间切换，运行的游戏不受影响）。从这个意义上说，OpenFlow将传统的物理固定的硬件定义互联网改造成为了动态可变的软件定义互联网（software defined networking）。而一个软件定义的可控的互联网，除了更加灵活以外，毫无疑问，通过恰当的控制算法，将大大提高网络自身的鲁棒性、运行效率以及安全性。

目前学术界OpenFlow主要是stanford、berkeley、MIT等牵头的研究组在推动，而业界据说包括Google在内的几大巨头已经纷纷参与其中，最新的版本1.0协议已经发布。牵头人Nick McKeown曾在Sigcomm08上做过专题的demo，后续这几年仍有不少的相关工作在高水平的会议上发表。国内据说清华大学已经有研究机构参与进去。

Nick McKeown这个人十分有意思（主页在<http://yuba.stanford.edu/~nickm>），现任stanford的AP，从他本人提供的简历就可以看出，Nick同学跟业界关系十分紧密，phd毕业两年就创办了公司，还参与了Cisco的项目，后来新公司卖给Cisco（Cisco这种模式很不错，有兴趣的同仁可以搜索过往案例）。笔者有幸在某次国际会议上碰到真人，给人感觉是十分的humorous且energetic的。Nick同学在推OpenFlow的时候明显十分重视跟业界结合，基本上是一边做，一边拉生产商的支持，很重视做demo，很早就在stanford的校园网中部署了OpenFlow，做的差不多了再提标准，再做宣传就事半功倍了。他的这种发展模式也十分为笔者所推崇。

最后的战役

OpenFlow的出现无疑给现有的交换市场带来了新的巨大的商机。网络行业发展到今天，垄断已经十

分的严重，许多年来，交换机制造商已经麻木于每天忙碌提高性能的目标，偶尔做点小工作，支持下出现的新的需求。而OpenFlow创造了一块前所未有的大蛋糕，能否抓住这一机遇，不夸张的说是重新瓜分市场的生死之战。目前Cisco、HP、Juniper、NEC等巨头已经纷纷推出了支持OpenFlow的交换设备，不仅有固网的，移动互联网领域也相关产品开始试水。从另外一个角度看，市场的重新瓜分，新需求的出现，也会给小规模的生产商带来一线生机，对于新出现的厂家来说，这也许是能争得一席之地最后的战役。

相关资料

官方网站见<http://www.openflowswitch.org>，目前有软件版本、netfpga版本可供使用，支持的交换设备相信很快就可以能够买到。