



SDN 与流量精细化运营

王素彬, 朱永庆

(中国电信股份有限公司广东研究院 广州 510630)

摘要:在分析网络流量疏导和优化多个案例的基础上,提出了一种通过 SDN 控制叠加网络来进行业务流量管控的思路和系统框架,该方案可以根据用户指定的策略对流量进行智能调节,可以被广泛应用于多种类型的 IP 网络,并可以有效利用成熟的协议和架构,支持现网平滑演进到 SDN。

关键词:软件定义网络;叠加网络;流量精细化运营

doi: 10.3969/j.issn.1000-0801.2014.011.025

SDN and Traffic Fine Operation

Wang Subin, Zhu Yongqing

(Guangdong Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China)

Abstract: On the basis of analyzing the real use cases of network traffic optimization, an idea and system architecture were presented, which abstracted the current network to the overlay network and used the SDN controller to control the overlay network for smart traffic steering and optimization according the policies defined by users. This solution can be widely applied in multiple different type of IP networks, and utilize the mature protocols to migrate the existed network to SDN smoothly with the benefits from SDN.

Key words: software defined networking, overlay network, traffic fine operation

1 SDN 的关键技术

SDN (software defined networking, 软件定义网络) 技术发展主要有以下 3 个方向。

(1) 叠加网络 (overlay network) 方案

叠加网络方案用于解决数据中心的虚拟化、自动化, 支持多租户的需求。叠加网络方案认为网络只要提供基本 IP 连通性, 不需要感知任何租户和虚拟化概念, 而是在服务器上直接发起虚拟化叠加在现有网络上, 这种方案通常也称为 IT 的多租户虚拟化方案, 主要的支持厂商包括 VMware 和微软。由于无法和网络联动, 叠加网络方案存在很多问题无法解决, 包括多播、QoS 以及服务器性能下降问题。

(2) 转控分离方案

将网络设备进行集中管控、转发控制分离, 简化转发设备。以 ONF 的 OpenFlow 协议为代表, 可以很好地支持开放编程、业务创新、业务自动化, 其代表厂商是 Big Switch。但是由于其技术标准还不成熟, 目前还难以满足运营商网络的复杂需求。

(3) 传统网管增强方案

以 Cisco 的 ONEPK 方案为代表, 这种方案能够解决业务自动化问题, 方案主要是开放设备的北向接口, 提供开放编程能力, 但是设备仍然保留原先的智能, 仅仅解决了业务自动部署问题, 但是对于快速业务创新、简化网络等方面都明显不足, 对比转控分离方案仍然存在差距。

虽然 SDN 的实现方案还存在争议, 但对于运营商来说, 庞大的现网存量设备如何向 SDN 方向平滑演进, 无疑



是一个更值得思考的问题。目前已有厂商对当前网络协议进行了扩展,使之能够在不影响现网的情况下,具备部分SDN的部分能力,例如针对IP网络的RR+方案和针对MPLS网络的PCE+方案。

RR+方案利用BGP路由反射器作为集中控制器,使用BGP作为控制器和路由器之间的控制协议,可以实现网络路径集中控制。控制器综合考虑网络拓扑、链路带宽、时延等属性以及所承载业务的质量要求,进行路径计算,使报文的转发路径能够不再以传统的最短路径方式转发,而是根据网络现状和业务需求,进行实时的灵活调整。

PCE+方案需要在MPLS网络中引入PCE服务器,路由器仍保留MPLS能力,但是将LSP计算托管到PCE服务器。PCE服务器根据网络实际情况进行统一的LSP路径计算,其计算结果相比传统的分布式计算方法更加智能。PCE服务器将计算结果下发至路由器,路由器根据PCE服务器的计算结果完成LSP的协商和建立。

RR+和PCE+方案可以使链路负载更均衡,使网络对业务更加敏感,并简化运维。虽然无法做到完全的转控分离,但是通过引入SDN集中控制理念,可以在不冲击现网的情况下引入SDN的部分优势。对于运营商来说,在SDN发展方向尚不明确的时候,通过部署SDN过渡技术对现网进行小规模的改造应该是更明智的做法。

2 运营商网络SDN经典应用场景：网络流量疏导与优化

管道是运营商的核心,但运营商不能只卖管道,管道要智能化,对流经管道的流量要可识别、可管理、可运营,所以越来越多的运营商将精细化流量运营作为其IP网络的核心诉求之一,这里面包含面向租户的和面向业务的两种类型。

(1) 面向租户(用户)的流量精细化运营

案例 1: 运营商希望在给企业用户提供专线接入的同时还能提供网络增值业务,利用管道优势增加管道的价值。

如图1所示,在这个案例里面,基于SDN控制器的解决方案可以识别企业用户的流量,负责将企业的业务流按需导入增值业务节点,实现业务链的功能,给企业用户提供业务自助订购、按需自动下放等差异化增值业务服务。

案例 2: 运营商希望给数据中心租户提供差异化服务,特别是VIP用户。对容易出现拥塞点(如数据中心出口)的流量按照不同租户的不同策略进行流量负载均衡。

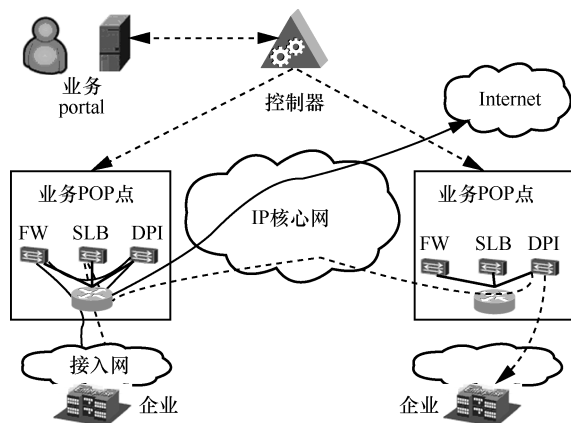
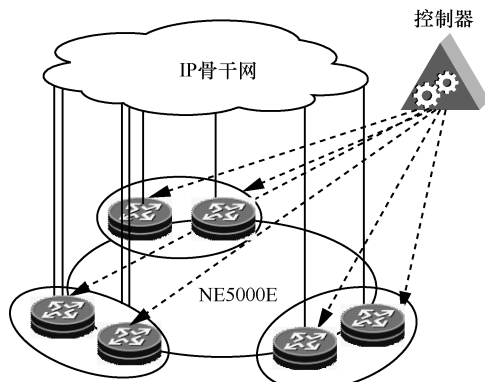


图1 流量精细化运营案例1: SPOP 增值业务链

如图2所示,在这个案例里面,基于SDN控制器的解决方案可以利用动态流量调优技术,识别出VIP用户,监控VIP用户的业务流量,对于超出预定门限的情况进行调整,为高价值租户提供差异化服务保证。



汇聚10个IDC机房, 出口宽带1 Tbit/s吞吐量

图2 流量精细化运营案例2: 数据中心出口差异化 QoS

案例 3: 运营商希望给企业用户提供 anti-DDoS 增值服务, 但又不希望全网每处都部署较昂贵的 anti-DDoS 清洗设备如图3所示, 在这个案例里面, 基于SDN控制器的解决方案可以将订购了 anti-DDoS 服务的用户业务流从攻击源头(A节点)疏导到最近的 anti-DDoS 清洗设备上(C节点), 而且会选择一条不拥塞的路径, 确保清洗过程中不会对其他业务造成影响。

(2) 面向业务的流量精细化运营

案例 4: 运营商希望按照业务系统的需要调整数据中心之间的流量, 对不同的业务提供不同的DCI服务。

在这个案例里面, 基于SDN控制器的解决方案可以识别不同的业务, 利用动态流量调优技术, 对DCI网络进行全局优化, 为不同的业务提供不同质量的链路。

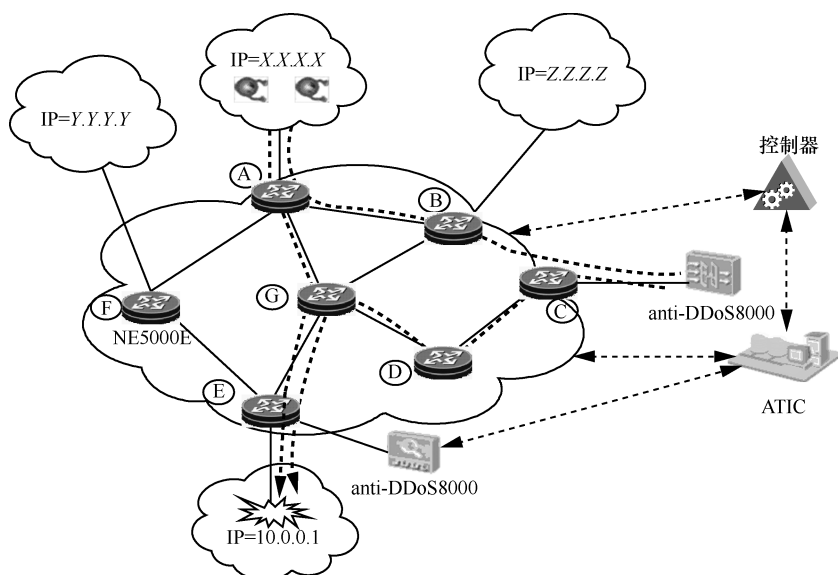


图3 流量精细化运营案例3:按需的 anti-DDoS 流量疏导

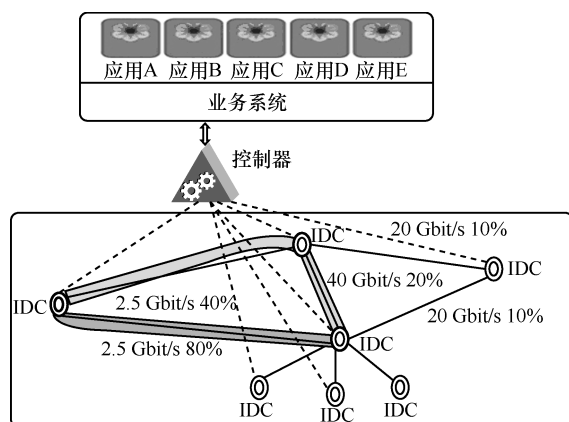


图4 流量精细化运营案例4:基于业务的 DCI 流量优化

不难看出,以上4个案例的核心共性需求是要求网络可以提供灵活智能的流量疏导能力和全局流量优化能力,这种流量疏导和优化能力是基于租户和业务的,而非原来基于网络的。

- 灵活智能的流量疏导能力,要求网络可以感知业务,按需疏导流量。
- 全局流量优化能力,要求网络存在一个智能设备,其拥有全局视图,可以全局监控并调整网络业务流量,达到全局最优。
- 基于租户和业务,要求网络具备开放性和可编程能力,可以根据用户千差万别的需求进行定制化处理。

而这些正是SDN的理念,也是SDN控制器有别于传统网络,最擅长的地方。

另外从业界来看,SDN目前最热的应用场景是在数据中心,而对于运营商网络,动态流量优化是目前唯一一个被较广泛接受的SDN应用场景。

总结,提供流量精细化运营的智能管道业务是运营商的核心诉求,而网络的流量疏导和优化能力是运营商使能智能管道的核心能力,SDN理念可以很好地适配这个能力,是SDN在运营商网络的最佳实践。

3 SDN 控制现网流量的主要思路

现在的IP网络是基于分布式路由计算的系统,每个路由节点会根据自己收到的路由信息计算从自己出发到各个节点的路径,然后控制流量的去向。路由协议需要负责在各个路由节点之间同步路由信息,确保全网的一致性和避免环路。如果SDN控制器直接修改其中某几条路由信息,可能会带来路由信息的不一致,甚至可能引起环路。

一个比较直接的想法就是抛弃掉现网的路由机制,采用一套全新的机制来控制路径,OpenFlow就是这个思路,通过全新定义的流表转发机制,控制器可以逐跳控制所有的业务流,达到流量疏导的目的。但这种方式也存在一些弊端:首先是对现网设备的升级改造,要求所有节点都要支持这个新的转发机制,包括硬件和软件都要升级;其次是对业务的适应性,不同的业务转发要求不同,要求OpenFlow协议需要不断地扩展以适配这些要求,协议的不完善将会导致业务功能的不完善以及互通存在问题,其实这里的许多问题在传统的协议里面已经得到了很好地



解决,但 OpenFlow 又要全新重来一遍,例如目前 OpenFlow 对运营商网络要求的 QoS 和 OAM 等能力普遍缺乏。

另一个可选的方案是保留现网的基础控制协议,而采用叠加网络的思路,控制器通过控制叠加网络达到对业务路径进行控制的目的。叠加网络的思路在传统 IP 网络上普遍存在的,它的思路可以追溯到 MPLS 的起源,MPLS 本质上就是在 IP 网络上叠加了一层以 LSP 为路径单位的 MPLS 网络,不论这个 MPLS 网络和 IP 网络如何映射,MPLS 协议控制的是这张 MPLS 网络,业务也是承载在 MPLS 网络的 LSP 路径上而非直接承载在 IP 路径上。

图 5 简要描述了 SDN 控制叠加网络的思路,这个叠加网络可以采用 MPLS、BGP/VPN、GRE、VXLAN、NVGRE、TRILL 等目前主流且比较成熟的技术,有效利用现有成熟协议和转发适配,而基础网络保持 IP/IGP 能力,可以让存量网络平滑演进到 SDN。SDN 系统采用统一的全局路径算法,但提供不同的南向接口来控制不同协议的叠加网络,这样可以适应于多种现有网络;叠加网络还可以让 SDN 聚焦于控制那些感知业务的节点,简化流量疏导和优化的复杂度。

不同的叠加网络所需要的 SDN 南向接口可以不同,比如控制 MPLS 网络,可以采用成熟的 PCEP;控制 BGP/VPN,可以采用成熟的 BGP;而控制 VXLAN,由于 VXLAN 没有定义控制层协议,则可以定义新的协议,当然也可以基于 OpenFlow 协议进行扩展。

另外,由于叠加网络隧道机制的不同,有以下两类控制模型。

- 逐跳隧道:MPLS、TRILL 等,适合于域内逐跳精确的

路径控制。

- 端到端隧道:BGP、GRE、VXLAN/NVGRE 等,适合于跨域端到端路径控制。

4 基于 SDN 的流量疏导与优化系统框架

如图 6 所示,一个完整的 SDN 流量调优系统应该包含如下 4 个核心部件。

- 流量分析器:负责收集现网的流量数据,包括对用户及对业务的识别,也包括对现网数据的汇聚、挖掘、分析等,并提供可视化报表给上层业务管理系统。
- 策略管理器:负责管理用户和业务策略,将合适的策略下放给控制器。
- 协同器:负责提供 API 及界面给上层业务管理系统,负责根据策略信息判定触发流量疏导和流量优化的动作,下放调优指令给控制器。
- 控制器:根据收集到的流量、策略和拓扑信息进行全局算路,并通过合适的南向接口将路径控制指令下到设备上去执行。

要完成一个端到端的流量调优,SDN 系统要采集以下 3 类信息。

- 流量信息:设备要支持 NetFlow、NetStream、SNMP 等机制负责上报带宽信息和流量信息,流量分析器负责对这些信息进行二次加工。
- 策略信息:这个信息可以是用户输入的策略,也可以是系统自定义的策略,这些策略要映射到对应的业务模型上去。

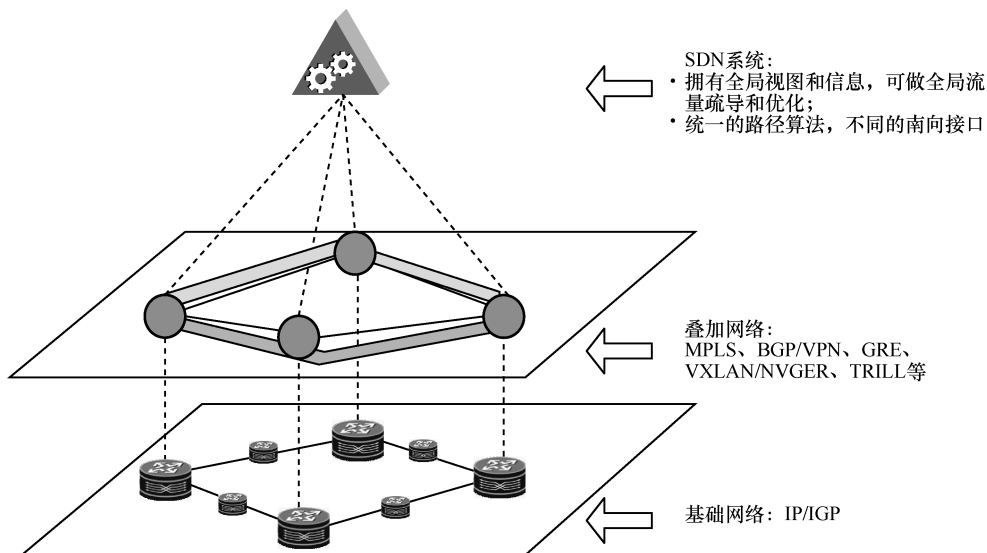


图 5 SDN 控制叠加网络的思路

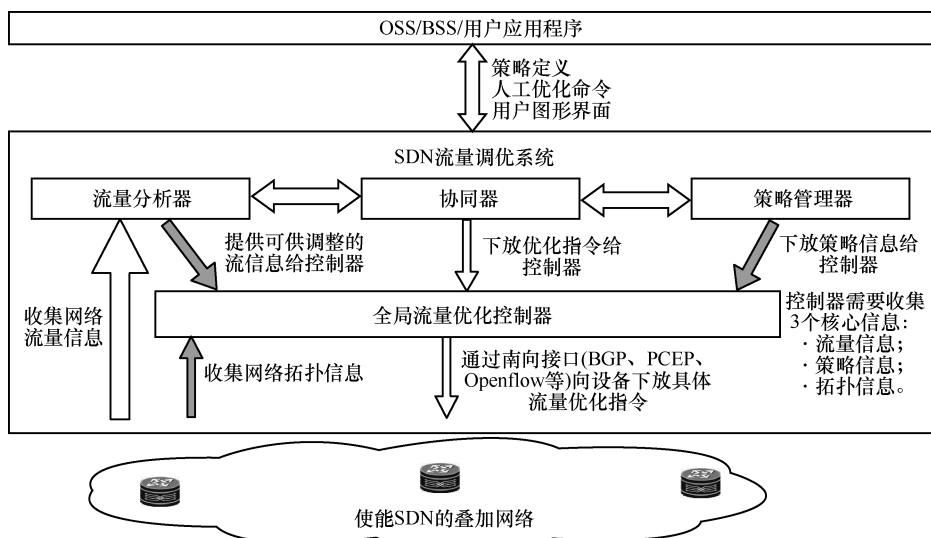


图6 基于SDN的流量疏导与优化系统框架

- 拓扑信息：拓扑信息可以由控制器直接去采集，控制器可以作为一个网元参与基础网络的拓扑协议，控制器也可以通过指定的拓扑协议收集现网拓扑信息。

对于路径计算，传统IP网络的流量工程在路由计算时仅考虑目的地址信息，不考虑源地址信息，更不考虑其他网络信息（如时延等），路由非业务想要的最佳。SDN系统的路径计算要支持多种不同的约束条件以及这些约束条件的组合，包括但不限于：

- 带宽利用率（上限、下限）；
- 链路开销或优先级（cost/metric 值）；
- 链路价格；
- 时延（含节点时延和链路时延）；
- 亲和属性（颜色）。

另外，策略管理器要支持对这些约束条件进行管理。

整个SDN流量调优的过程是一个闭环的过程：首先是设备上报流量和拓扑信息给SDN系统，然后是流量分析器做流量分析，挖掘出可供调节的流，接着是协同器根据流量信息和策略管理器提供的策略信息进行判定，将调优指令下给控制器，控制器根据流量、策略、拓扑等信息进行路径计算，为业务生成新的路径，再将这些新的路径通过南向协议下到设备上去执行，设备完成路径调整后上报新的流量和拓扑信息给SDN系统，从而完成闭环过程。

值得注意的是，虽然下面的网络可以是多种不同的叠加网络，但流量采集、流量分析、拓扑管理、策略管理、业务协同、路径计算等都是类似的，可以采用同一套框架来实

现。主要不同的是在南向接口以及为了适应不同叠加网络的一些特殊处理，如控制VXLAN需要有VTEP广播域的管理和VXLAN隧道的管理等，再比如控制BGP网络需要有对BGP邻居关系的管理等。

图7为SDN流量调优系统与SDN标准模型的对比，可以看出SDN流量调优系统吸取了SDN的两大核心理念（开放可编程、集中控制），并融合了运营商网络现有的成熟机制，更适合于运营商网络，让运营商在享受了SDN带来的好处的同时，又不必对存量网络进行革命性的改变，是SDN在运营商网络较为理想的演进方案。

5 实例：SDN系统应用于基于BGP的流量调优

下面举一个例子说明SDN流量调优系统的工作原理。

如图8所示，对于IGW（国际关口局）和DC（数据中心）出口等场景，运营商在流量均衡上面临如下挑战：

- 出口带宽利用率不均，造成扩容频繁，流量调整频繁；
- 流量不感知链路成本，造成网间结算费用增多；
- 流量不感知链路SLA，造成VIP用户体验差，投诉多。

由于AS（自治系统）的不同，当前的做法只能更多基于人工的BGP策略调整，这种手工调整的方式不能实时调整、耗时长、配置和维护复杂。

采用SDN流量调优系统可以帮助运营商进行自动化的流量调整，降低手工维护的工作量，最大化IGP/DCI带宽利用率，降低网间结算费用，并给不同的用户/业务提供不同的带宽/QoS保证。

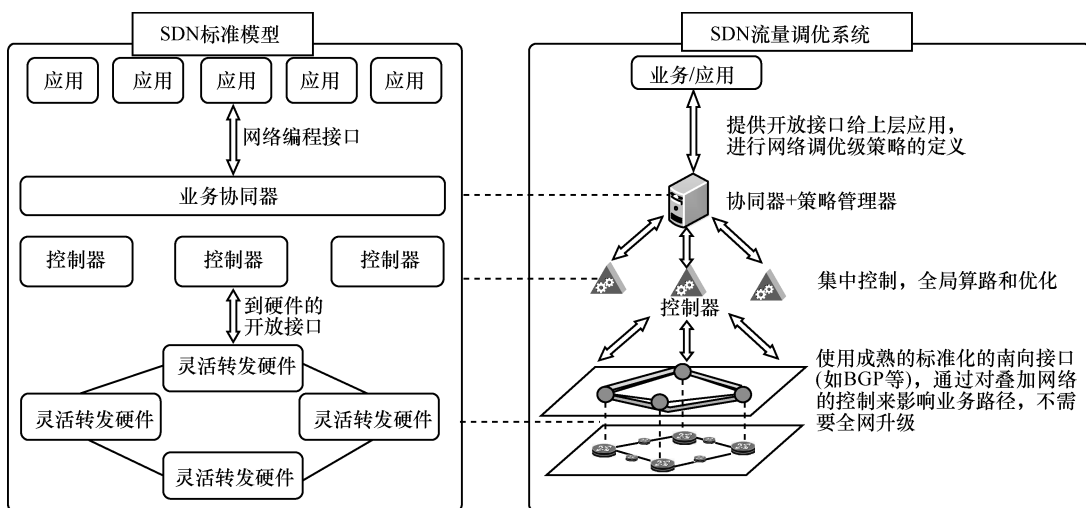


图7 SDN 流量调优系统与 SDN 标准模型对比

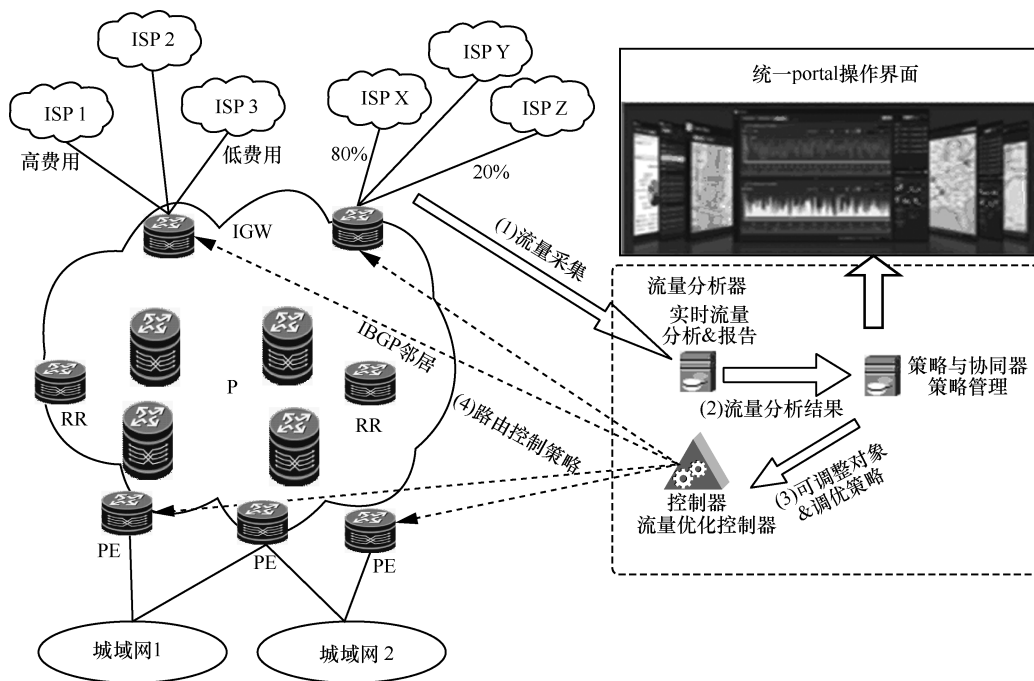


图8 SDN 系统应用于 IGW(国际关口局)和 DC(数据中心)出口流量调优

流量分析器利用 SNMP 和 NetStream(NetFlow)分别采集链路带宽利用率和聚合流量信息。

(1)基于 SNMP 收集链路带宽利用率

- 入方向流量；
- 出方向流量。

(2)基于 NetStream 收集聚合流

- 基于目的 IP 地址的聚合流；
- 基于目的 AS 的聚合流。

针对如图 9 所示的组网结构，流量分析器分别在

IGW1 和 IGW2 上按链路采集链路信息，并对各个流进行排序，识别出各个链路上的 TOPN 大象流(流量排名靠前的流)，将这些流作为“可调整的对象”，如图 10 所示。

策略管理器设置调优算法的约束条件举例如下：

- 带宽利用率上限，如小于 80%；
- 带宽利用率偏差门限，如小于 10%；
- 链路价格因子，如小于链路 1 成本= $A+B(X-X_0)$ ；
- 流量调控方向，如小于入方向或出方向；
- 流量统计分析间隔时间，如小于 5 min。

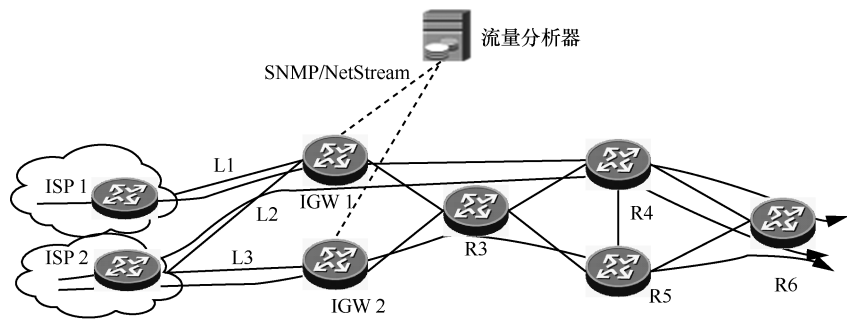


图 9 基于 BGP 流量调优的示例组网

链路统计报表(举例)				流统计报表(举例)			
路由器	链路	入方向	出方向	入方向流量排序		出方向流量排序	
IGW 1	L1	80%	30%	flow 1	400 Mbit/s	flow 1	200 Mbit/s
	L2	45%	30%	flow 2	350 Mbit/s	flow 2	90 Mbit/s
IGW 2	L3	20%	10%	flow 3	50 Mbit/s	flow 3	10 Mbit/s
				flow 4	250 Mbit/s	flow 4	200 Mbit/s
				flow 5	200 Mbit/s	flow 5	100 Mbit/s
				flow 6	100 Mbit/s	flow 6	50 Mbit/s
				flow 7	50 Mbit/s	flow 7	40 Mbit/s
				flow 8	50 Mbit/s	flow 8	10 Mbit/s

图 10 流量排名

策略管理器还需要设置调优目标。

- (1)选项 1:带宽利用率均衡。带宽利用率小于预设的门限,同时带宽利用的偏差小于预设的门限。
- (2)选项 2:价格最优。优选价格最小的链路,同时这个链路上的带宽利用率小于预设的门限。

当然,也可以由用户自定义目标。

协同器负责提供 API 及界面给上层业务管理系统,支持以下方式。

- (1)手工拖拽(针对指定流,直观调节),如图 11 所示。
- (2)调整目标:流 x (200 Mbit/s)从链路 A 调至链路 B,如图 12 所示。

调整前:链路 A 流量 800 Mbit/s,链路 B 流量300 Mbit/s。
调整后:链路 A 流量 600 Mbit/s,链路 B 流量 500 Mbit/s。

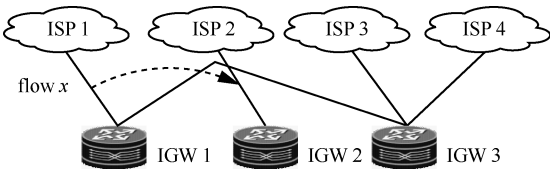


图 11 手工拖拽

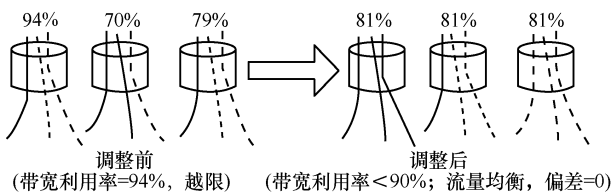


图 12 调整前后对比

- (3)自动优化(设置触发因子,如带宽利用率越限),见表 1。

表 1 自动优化	
优化目标	出方向带宽利用率均衡
节点	IGW1, IGW2..., IGWn
待调整对象	TOP N 大客流
门限	带宽利用率<90%(触发门限) 利用率偏差<10%

流量分析器、策略管理器要把可调整的流的信息和调优策略信息通过 Netconf 等协议下发给控制器。协同器在触发调整门限后,将调整命令通过 Netconf 下发给控制器。
控制器需维护网络拓扑和 BGP 关系,将对端 AS 抽象为网络拓扑上的一个节点,如图 13 所示。

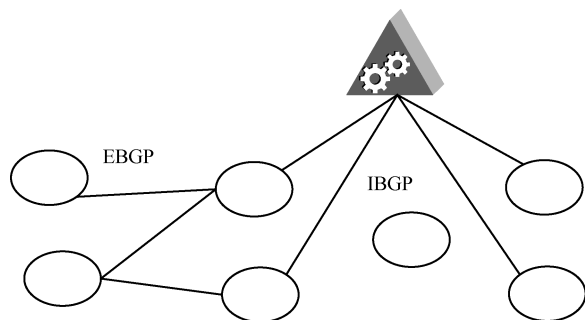


图 13 网络拓扑抽象图

按照拓扑和 BGP 关系自动生成 EBGp export 策略:设置不同的 community,对应一大一小 MED 和一长一短 AS-path。

例如图 14 所示的组网,分别在 IGW1 和 IGW2 上生成如下预配置。

(1)IGW1 给 ISP1 的 EBGp export 策略

If match community1, MED=大;
If match community2, MED=小;
If match community3, AS-path=加长;
If match community4, AS-path=不变。

(2)IGW1 给 ISP2 的 EBGp export 策略

If match community5, MED=大;
If match community6, MED=小;
If match community7, AS-path=加长;
If match community8, AS-path=不变。

(3)IGW2 给 ISP2 的 EBGp export 策略

If match community1, MED=大;
If match community2, MED=小;
If match community3, AS-path=加长;
If match community4, AS-path=不变。

然后控制器就可以根据其获得的流量、策略和拓扑信息进行带多重约束条件的全局路径计算,将 flow2 从 L1 调整到 L3,把 flow8 从 L3 调整到 L1,从而满足用户指定的约束条件:带宽<70%,偏差<10%,如图 15 所示。

下面讨论如何把这些具体的路径调整指令下放到设备上执行。

针对要调整流,控制器给不同 IGW 下发不同的路由及对应的团体属性,例如控制器生成 IGW1 设备的 BGP 路由及属性:

IP-prefix 2(flow 2), community3, community5

IP-prefix 8(flow 8), community7, community4

生成 IGW2 设备的 BGP 路由及属性:

IP-prefix 2(flow 2), community2, community4

IP-prefix 8(flow 8), community1, community3

这些携带了团体属性的 BGP 路由通过 BGP 下发到设备 (IGW1 和 IGW2) 上,命中设备上预先配置的 EBGp export 策略,从而执行相应的动作,如变更 MED 值或变更 AS-path 长度等,向对端 AS 发布更新的 EBGp 路由,从而

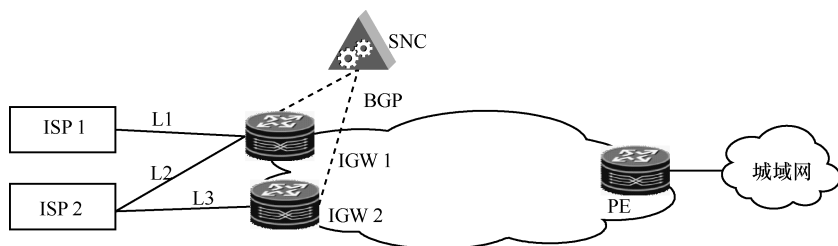


图 14 BGP community 预配置示意

实时流量报表&约束定义				优化后的流量分布		
L1	80% (越限)	flow 1	400 Mbit/s	49%	flow 1	400 Mbit/s
		flow 2	350 Mbit/s		flow 3	50 Mbit/s
		flow 3	50 Mbit/s		flow 8	40 Mbit/s
L3	20%	flow 6	100 Mbit/s	51%	flow 2	350 Mbit/s
		flow 7	60 Mbit/s		flow 6	100 Mbit/s
		flow 8	40 Mbit/s		flow 7	60 Mbit/s
L2	45%	flow 9	50 Mbit/s	45%	flow 9	50 Mbit/s

要调整流

均衡: 满足带宽<70%, 偏差<10%

图 15 流量调优示意

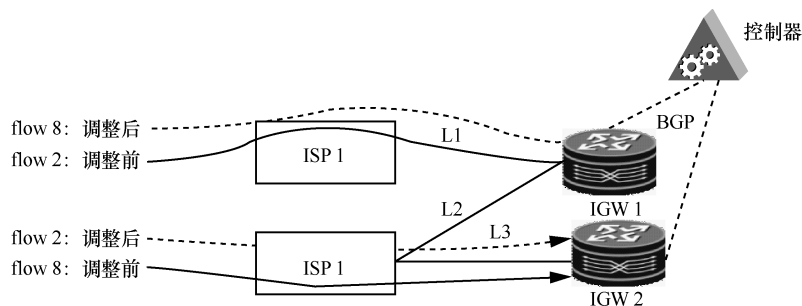


图 16 调整效果示意

达到控制流量路径的目的。

最后的调整效果如图 16 所示(以调整入口流量为例)。

值得注意的是,整个调整过程虽然是自动的,但也可以支持确认机制,即调整执行前由用户进行手工确认。

图 17 为对 BGP 网络流量调优系统的总结。

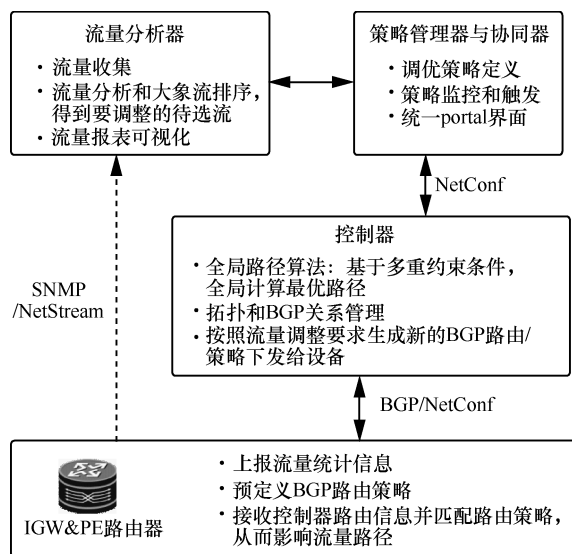


图 17 BGP 网络流量调优系统

6 结束语

SDN 流量疏导与优化系统是 SDN 在运营商网络的最佳实践,它利用 SDN 集中控制与开放可编程的理念,通过对叠加网络的控制,实现对业务流量的控制,帮助运营商实现流量的精细化运营,并将现网平滑演进到 SDN 架构。

这个方案的灵魂在于其策略和算法的开放性以及对现网的兼容性,尤其是采用对叠加网络的控制给这个方案带来了无穷的想象空间,例如只要是 BGP 适用的场合,就有机会应用基于 BGP 的控制方案,只要有

MPLS 的地方,就可以考虑应用基于 MPLS 的控制方案。通过集中控制和开放可编程,给传统网络增加智能化能力。未来可以预期,SDN 将在更多的如下场合中,提供流量疏导与优化方案。

- 移动承载的 Seamless MPLS 网络;
- 跨越多张 IP 核心网络的端到端业务;
- 数据中心内加数据中心间的端到端网络;
- 城域以太二层网络或基于 GRE 的三层网络;
- 固网 POP;
- 移动 Gi-LAN 等。

参考文献

- 1 McKeown N, Anderson T, Balakrishnan H, et al. OpenFlow: enabling innovation in campus networks. ACM Communications Review, 2008(4)
- 2 中国电信集团科学技术委员会. 中国电信 IP 骨干网络架构研究, 2012

[作者简介]



王素彬,女,中国电信股份有限公司广东研究院工程师,从事 IP 网络技术和下一代互联网研究工作。



朱永庆,男,中国电信股份有限公司广东研究院工程师,主要研究方向为超宽带城域网技术研究。

(收稿日期:2014-05-23)