



专题：SDN/NFV

重构广域网关键技术

郭晓军, 万晓兰

(新华三技术有限公司, 北京 100085)

摘要: 传统广域网存在诸多问题, 变革势在必行, 重构广域网意义重大, 从而引出软件定义广域网(SD-WAN)重构方式。综合现阶段多种技术, 对 SD-WAN 定义、架构到开放平台、控制平面协议、南向协议等方面进行分析 and 探讨, 如何选择具体技术需要结合技术本身成熟度和应用实践来决定。

关键词: 重构; 广域网; 软件定义广域网

中图分类号: TN915.43

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017099

Key technology of reconstructing WAN

GUO Xiaojun, WAN Xiaolan

New H3C, Beijing 100085, China

Abstract: There are many problems on traditional WAN, change is imperative, and it is very important to reconstruct WAN, which leads to software defined-wide area network (SD-WAN) reconstructing style. Among the current technology, the definition of SD-WAN, SD-WAN architecture, ODL open platform, protocol of control platform and southern protocol were analyzed and discussed. How to choose specific technology needs to combine technology maturity and application practice to decide.

Key words: reconstruct, WAN, SD-WAN

1 引言

长期以来, 传统广域网(wide area network, WAN)存在着部署周期长、管理复杂度高、业务适应性差、系统不开放等问题, 这些问题一直困扰着网络运维人员, 也制约着企业的 IT 发展。尤其近年来随着云计算、在线视频、移动互联网等新兴应用模式的快速发展, 从根本上改变了广域网的流量模型, 网络需要承载的应用种类越来越多, 流量也变得更加复杂多变, 这对广域网的承

载能力、业务适应性、调度灵活性、可扩展性、可靠性等都提出了新的要求, 这些新的需求使得广域网面临的问题更加突出和复杂, 这些问题在现有的网络架构上很难根本解决, 因此, 急需对传统的广域网进行重构和变革, 构建一个架构开放、灵活编程、易于运维的新一代广域网, 承载日益丰富的应用流量, 最终实现应用按需驱动网络、网络动态适应应用, 满足用户的应用需求。

重构广域网有很多技术和手段, 可以说“条条大路通罗马”, 具体选择哪条路, 就要结合当前

的技术成熟度和应用实践来确定,本文着重探讨重构广域网的一些关键技术。

2 广域网重构的核心思想

SDN (software defined networking, 软件定义网络) 作为目前最热门的思想和技术,已经在数据中心、园区网等领域得到广泛应用和规模部署,使网络具备灵活编程、动态感知及自动编排等能力,从而快速满足不断变化的应用需求。虽然广域网和数据中心等网络有很多不同,但在满足用户应用需求上是相同的,因此,将 SDN 引入广域网领域中已经成为业界的普遍共识,也就是本文要讨论的 SD-WAN (software defined-wide area network, 软件定义广域网)。

2.1 SD-WAN 的定义

SD-WAN 作为重构广域网的指导思想和核心技术,不仅能够提供与 SDN 相同的优势,而且通过虚拟化资源、自动化部署、可视化管理及智能化调度等手段,加速服务交付,提高广域网性能及可用性,简化运维管理,降低总体成本。从定义可以看到,SD-WAN 具备以下特点。

- 自动化: 将用户定义的需求和策略进行解析,自动转换为合适的网络模型,分配相应的网络资源,最终翻译成设备命令,下发到设备上,实现业务的自动部署,缩短上线周期,促进网络从功能向服务转型,产生增值。
- 全局视角: 通过集中的控制平台,对网络中的设备和链路进行抽象和虚拟化,形成统一的网络资源池,从全局的视角来进行管理和控制,优化流量分布,保障业务质量,提高网络利用率。
- 智能控制: 以用户应用为中心,根据当前的网络运行状况,从全网视角为应用动态选择满足带宽和质量需求的最优路径,并根据网络和流量变化动态进行调整,确保

应用的服务质量,持续提供服务。

- 可视化: 通过丰富的采集和探测手段,收集网络运行状态和应用流量等数据,进行分析和整合,从设备、链路、应用等多个层面和维度进行可视化呈现,进一步简化广域网的运维管理。
- 开放: 采用 SDN 架构,使各组件解耦,从设备到控制器再到 App 提供多层次的、不同抽象度的 API,实现整个网络的可定义、可编程,差异化定制网络服务,促进 IT 和 CT 融合,快速适应业务需求,带来新的增值。

2.2 SD-WAN 架构下的广域网

传统广域网是一个复杂封闭、被动承载的网络,网络和应用是相对割裂的,它们之间的关联是通过管理员人工关联起来的,这种关联是静态的,无法根据应用需求的变化而变化,SD-WAN 采用集中的控制平台,通过南向协议和北向接口将用户的应用系统和网络有机地联系起来,一方面让网络更加主动、实时感知应用,适应应用的需求变化;另一方面也让应用可以很方便、灵活地使用网络服务,驱动网络为应用服务,让整个广域网从被动网络变为主动网络,具备开放的网络架构、可视的状态呈现、全局的控制调度、简化的运维部署、灵活的编程定制等能力,动态保障应用的服务质量和用户体验。SD-WAN 架构如图 1 所示。

3 开放架构的支撑平台

相对传统广域网,在 SD-WAN 架构的广域网中,集中的控制平台是一个非常重要的组件,这个平台直接影响整个网络的开放性和灵活度,因此,对重构一个广域网来说,选择一个好的控制平台至关重要。

3.1 开放的 ODL 平台

ODL (OpenDaylight) 是 Linux 基金会旗下的

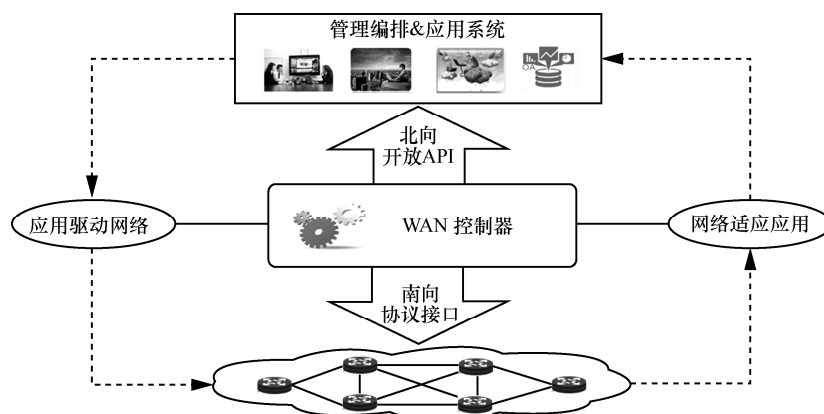


图1 SD-WAN 架构

一个开源控制器合作项目，致力于制定一个共同的、开放的 SDN 平台为开发者使用，促进和建立商业产品和技术的发展，加速 SDN 产业的创新。目前已经有包括 AT&T、NEC、Intel（英特尔）、Cisco（思科）、HP（惠普）、H3C（新华三）、华为等在内的众多设备商和运营商参与其中，每个成员可使用 ODL 提供的插件，使其产品、服务增强，为用户带来附加价值。

ODL 项目于 2013 年 2 月启动，并于 2014 年 2 月发布第一个正式版本，之后每隔 8 个月发布一个正式版本，目前最为成熟稳定的是第 4 个版本——铍版本，它能够为用户提供和企业提供自动化服务交付、网络资源优化、云计算、NFV 以及区域网络的自动化、可见性和可控制相关的解决方案。

ODL 作为参与度最广的、领先的开源控制器平台，已经在数据中心网络广泛应用，现在已逐步向广域网领域延伸和普及，主流 IT 网络提供商争先恐后地将 ODL 引入广域网解决方案中，作为构建开放广域网的集中控制平台，充分利用和发挥开源的优势和力量，以开源、开放的 App 开发模式，满足用户不同场景下的应用需求。

3.2 灵活的 App 开发模式

进行 ODL 的 App 开发，最为重要的是把 Yang 与 MD-SAL（model-driven service abstraction layer，MD-SAL）这两个关键技术点相融合。

Yang（RFC 6020）是用于网络配置协议

（NETCONF）的建模语言，一个 Yang 模型定义了一个数据的层次结构，并把其模型化为一棵树，树中每个节点都有名称，且要么有一个值要么有一个子节点集。Yang 还可以用于基于网络配置协议（NETCONF）的操作，包括 RPC（remote procedure call，远程过程调用）和 notification（通告）等。铍版本依赖关系中，位于最底层的 YangTools 模块几乎被所有的模块所依赖，是整个 ODL 项目的核心。YangTools 通过 Yang 模式来规范数据格式，大大简化了开发者的代码编写。配合生成器可以直接生成 Java 代码，保证了调用模块间数据格式一致，并且可以直接通过 Java 对象传递。

ODL 要加强基于 MD-SAL 的架构，MD-SAL 采用模型驱动的方式对业务进行抽象，为南向和北向 API 的统一以及在控制器内部对不同业务功能组件的数据结构定义提供了便利。为了完成这一功能，MD-SAL 使用 Yang 作为其建模语言来对业务和数据进行抽象，并引入了消费者和提供者概念，SAL 中的 API 由模型提供并自动生成。SAL 层并不定义具体的模型。用户可以自定义新的模型，并添加到 MD-SAL 中即可。MD-SAL 不感知插件之间的具体数据信息。MD-SAL 中提供 3 种核心服务：用于数据库存储的数据存储（data store）服务，用于模块间调用的 RPC/业务路由（service routing）服务以及模块间的通知订阅和发布服务。

基于 Yang 模型的 App 开发过程如图 2 所示。由图 2 可知,在定义好 Yang 模型后通过 YangTools 生成 Java 接口和类,开发者可以使用模型定义好的数据结构和代码框架,编写插件代码,再由 Maven 工具生成 OSGI (open service gateway initiative, 开放业务网关计划) 框架下的 API bundle 和插件 bundle,最后打包成一个特性部署到控制器中。

对于 ODL 的 App 开发而言,通常分为 3 个层次。

(1) 基于 RESTful API 的上层应用开发

该方式的关注点更多地集中在 Web 可视化界

面的开发上。此外,针对已有非 OSGI 框架、非 Java 开发语言、使用独立数据库的北向编排系统,如果需要与 ODL 进行对接,也采用该方式。

(2) 基于 MD-SAL 的北向 App 插件开发

基于 MD-SAL 的北向 App 插件开发如图 3 所示。在 MD-SAL 内部进行开发,一方面可以借助 MD-SAL 提供的模块间高效和灵活的访问方式,对 ODL 现有网络特性进行特色整合;另一方面,通过 Yang 模型的建模,提高开发效率。另外,合理地使用 MD-SAL 自带的数据存储内存数据库,可以很容易地实现持久化数据和非持久化数据的存储,甚至开发出分层次的集群特性。

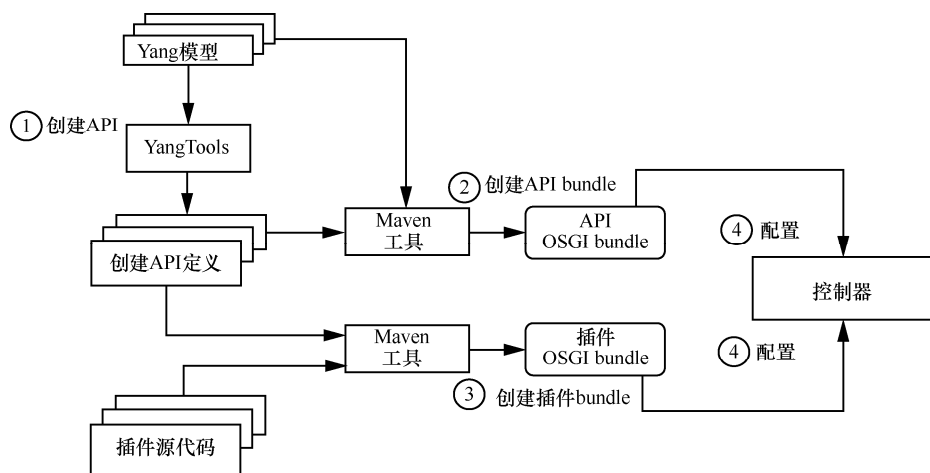


图 2 基于 Yang 模型的 App 开发过程

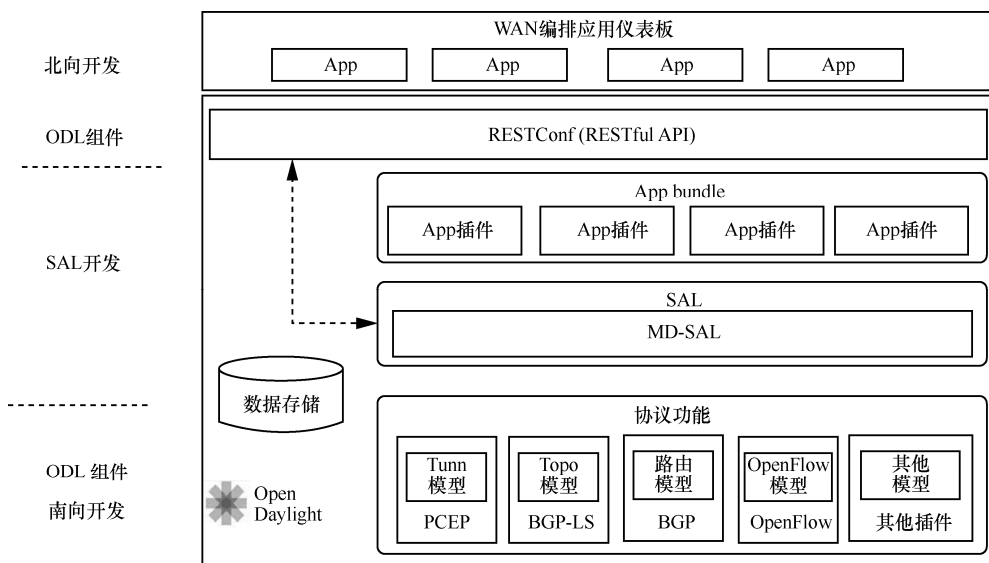


图 3 基于 MD-SAL 的北向 App 插件开发



(3) 基于 MD-SAL 的南向协议插件开发

由于 ODL 支持丰富的南向协议，目前南向插件的开发主要集中在协议的增强以及北向 App 插件或外部 App 的适配上，通常并不是 App 开发的重点。

综合以上 3 种方式来进行广域网 App 的开发，在控制器内部，开发北向 App 插件（例如开发拓扑、流量调度、协议配置等），用于整合南向协议和现有北向服务，通过 MD-SAL 为插件自动生成 RESTConf API 为上层 WAN 编排模块进行调用。另外，通过 Yang 扩展，在不需要对 ODL 现有模块进行修改的前提下，为南向协议模块编写扩展插件，增强其协议功能。

ODL 作为一个开源开放的平台，可以很灵活地进行广域网 App 的开发和增强，具有良好的扩展性，也为构建一个开放的广域网提供了坚实的支撑。

4 灵活高效的控制平面

智能控制、灵活调度是 SD-WAN 的一个重要特征和关键能力，这样才能为应用报文选择最优的网络路径，更好地保障应用的网络质量和用户体验，因此，选择一个灵活高效的控制平面至关重要。

分段路由（segment routing，SR）作为一种源路由协议，为 SD-WAN 提供了更加灵活简单的路径控制方式，逐渐成为 SD-WAN 集中的控制平面，实现全网的流量调度和路径优化，保障关键业务质量，均衡流量分布，提高专线利用率，降低线路成本。

4.1 SR 介绍

SR 是一种源路由架构协议，由源节点来选择或指定路径，并将路径转换成一个有序的分段列表封装到报文头中，核心设备只需要根据报文头中的路径进行转发。SR 是对现有源路由和流量工程的完善，只需要在边缘设备进行路径选择，消除了核心设备对路径状态的维护，简化了广域网的设计和管理。

简单地说，SR 就是对现有控制平面进行简化和直接复用已有的转发平面。

- 简化控制平面：对现有的控制平面进行简化，如在 MPLS（multi-protocol label switching，多协议标签交换）网络中，不再需要部署复杂的 LDP（label distribution protocol，标签分配协议）或 RSVP-TE（reserved for tunnel resource，隧道资源预留）协议，只需要设备通过 IGP（interior gateway protocol，内部网关协议）路由协议的 SR 扩展来实现标签分发和同步，或者由控制器统一负责 SR 标签的分配，并下发和同步给设备。
- 复用转发平面：直接复用已有的 MPLS 和 IPv6 转发平面，网络设备不做改动或者进行小的修改就可以支持对 SR 的转发，如在 MPLS 网络中，segment 就是 MPLS 标签，路径就是标签栈；在 IPv6 网络中，segment 就是 IPv6 地址，路径就是封装在一个路由扩展头中的 IPv6 地址列表。

4.2 SR MPLS 转发平面

在 MPLS 转发平面中，利用标签栈作为路径，报文格式如图 4 所示。



图 4 标签栈作为路径的报文格式

这里的标签与普通 MPLS 的标签格式完全相同。SR 标签分为两种：节点标签（prefix/node segment）和邻接标签（adjacent segment）。

节点标签是为设备本身分配的标签，一个设备即一个节点，标签全局唯一。每个设备有各自支持的 MPLS 标签范围，支持 SR 的设备需预留一段标签段作为全局节点标签 SRGB（segment routing global block）。设备通过 IGP 扩展通告

SRGB 的标签段范围以及自己的全局索引, 每个设备节点对应的节点标签即 SRGB+索引。

节点标签寻址示意如图 5 所示, R6 节点通过 IGP 扩展发布了自己的索引为 6, SRGB 从 7000 开始, 则 R6 的节点标签为 7006。因此当网内其他设备收到栈顶标签为 7006 的报文时, 会直接查找一条到 R6 的最短路由并将报文发送给 R6。

如从 R1 开始到 R6 的报文, 处理流程如图 6 所示。

- 首节点 (R1): 根据 SDN 路径计算结果, 确定报文需要经过 R6, 加入节点标签栈 7006。
- 转发节点 (R2): 与普通 MPLS 报文处理一致, 进行标签交换处理, 标签仍然为 7006。

- 倒数第二跳 (R3): 如 IGP 通告 SR 信息时, 通告需要进行倒数第二跳弹出, 则在此节点进行标签弹出。
- 尾节点 (R6): 按照普通报文正常流程进行转发。

邻接标签表示设备上某条链路的单跳路径, 标签仅在设备本地有效。每个设备向与自己一跳相邻的设备通过 IGP 扩展通告邻接标签, 或者通过 SDN 控制器直接为 SR 域内的每条链路进行标签分配。邻接标签转发示意如图 7 所示。

- 首节点 (R1): 根据 SDN 路径计算结果, 确定报文路径为图 7 中箭头所指向的线。根据邻居标签信息在 R1 处打上 [204,405,506] 标签栈。
- 转发节点 (R2、R4、R5): 弹出栈顶标签,

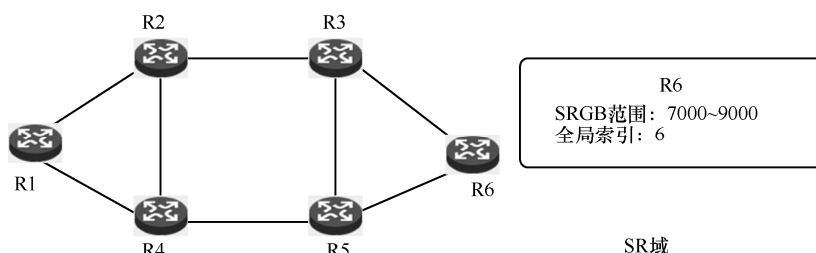


图5 节点标签寻址示意

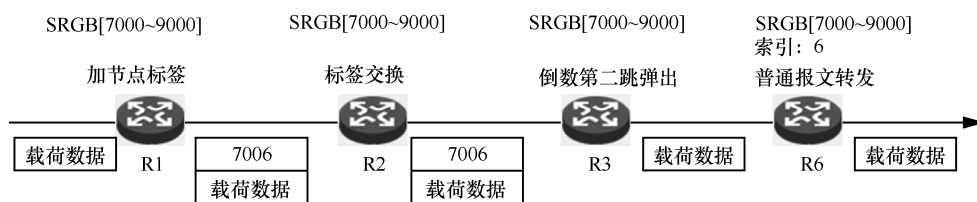


图6 报文转发示意

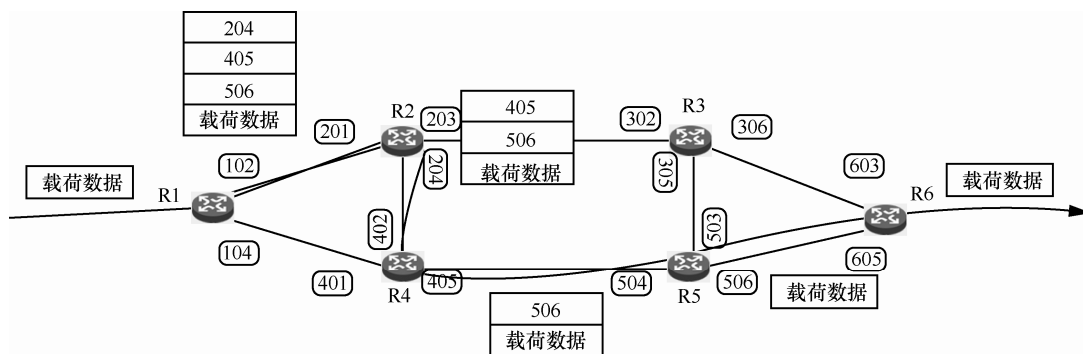


图7 邻接标签转发示意



根据此标签选择出口链路进行转发。

- 尾节点 (R6): 按照普通报文正常流程进行转发。

4.3 SR IPv6 转发平面

IPv6 转发平面中, IPv6 地址即 segment, 通过 IPv6 地址序列来表示报文转发路径。基于 IPv6 报文的原有路由由扩展头定义了新的 SRH (segment routing header), 用于记录 segment 的相关信息。具体数据报文格式如图 8 所示。通过这样只会增加扩展头的方式, 可与原有 IPv6 的转发平面平滑融合。



图 8 segment 报文 IPv6 格式

如从 R1 到 R6 的报文希望经过箭头指向的线所指示的路径, 在 IPv6 转发平面下的处理流程如图 9 所示。

- 首节点 (R1): 根据 SDN 路径计算结果, 确定报文路径为 R2、R4、R6, 增加 SRH, 用 IPv6 地址栈的方式指明中间需经过的路径为 R2、R4、R6, 此时需将 IPv6 目的地址改为下一个待 SR 处理的节点地址: R2, 同时将路径中指示下一个待处理 segment 为 R4。

- SR 路径涉及转发节点 (R2、R4): 如果发现 IPv6 头中目的地是本节点, 则查看 SRH, 如果仍有下一个待处理 segment, 则将目的地址改为 SR 路径的下一个节点, 同时更新待处理 segment 指针。
- 非路径涉及转发节点 (R5): 直接根据 IPv6 报文头中的目的地址进行 IPv6 报文转发, 不修改报文内容。
- 尾节点: 去掉 SRH, 按照普通 IPv6 报文进行转发。

从上述流程可以看出, 对于非 SR 路径涉及的转发节点, 完全无需支持 SR, 仅支持标准的 IPv6 转发即可。采用这种方式, 对网内不支持 SR 的设备可以进行无缝兼容。

4.4 SR 的优势

SR 有很多优势, 详细介绍如下。

首先, 面向 SDN 架构设计的协议, 融合了设备自主转发和集中编程控制的优势, 能够更好地实现应用驱动的网络。同时, 可以天然支持传统网络和 SDN, 兼容现有设备, 保障网络平滑演进。

其次, 简化设备控制平面, 减少路由协议数量, 简化运维管理, 降低运营成本; 标签转发表简单, 容易扩展, 规模也很小, 一台设备上维护的转发表数为 $N+A$, 其中 N 为节点标签数量, 一般为全网节点数量; A 为邻接标签数据, 一般为设备接口数量, 而传统 MPLS 网络的转发表数则为 N^2 。

最后, 支持广泛的部署场景, 包括骨干网、

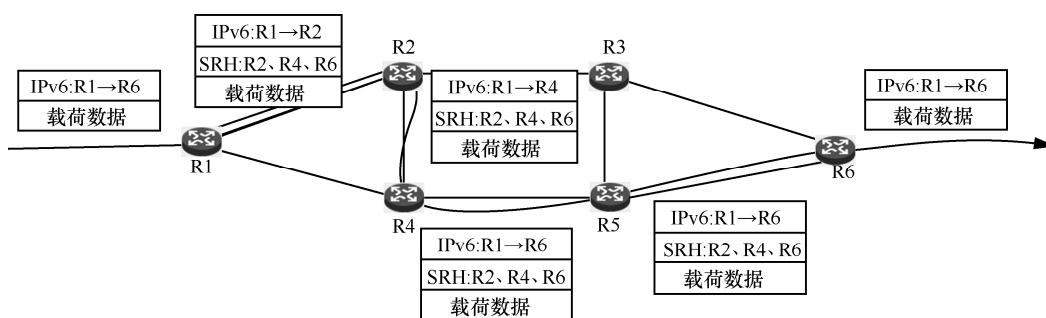


图 9 IPv6 转发示意

DCI 网络,可以同时支持 MPLS 和 IPv6 网络。

SR 作为一个关键技术,为广域网重新构建了一个集中的控制平面,在灵活调度广域网资源的同时,保障了传统广域网到 SD-WAN 的平滑演进。

5 多样的南向协议

广域网重构,其实就是 SD-WAN 的应用,其中最重要的就是控制平台如何和网络设备打交道,即南向接口协议,简单来说,包括采集和控制两类协议。采集就是控制平台通过协议接口来获取网络拓扑、运行状态、网络质量及应用流量等信息,为网络可视化、智能控制提供大数据支撑;控制就是控制平台通过分析采集上来的数据,并结合用户策略和应用需求,动态转换成网络指令,通过协议接口来实时控制网络的转发行为。

5.1 网络可视化/控制的基础——BGP-LS

传统广域网很大的问题就是网络不够透明、可视化程度不够,这导致管理人员很难管理整个网络,出现问题难以定位,运维压力很大。广域网重构的目标之一就是增强整个网络的可视化,使管理人员可以很方便地看到整个网络的拓扑结构、链路状态、链路质量及流量路径等,这其中用到的最基础的就是拓扑收集协议 BGP-LS (border gateway protocol-link state)。BGP-LS 可动态收集网络拓扑、设备信息及链路状态等数据,作为网络可视化、流量调度的基础,目前,BGP-LS 已成为广域网控制器的主流南向接口协议之一。

BGP-LS 是在 BGP 的基础上进行扩展,如图 10 所示,用来发布网络中设备节点信息、链路属性(如带宽、开销等)、链路状态及拓扑信息等。随着应用的发展,会有更多的扩展,如 SR 标签等,都会通过 BGP-LS 进行携带和发布,以实时收集更多的网络信息,支撑网络的可视化和精细调优。

域内信息收集如图 11 所示,BGP-LS 以域为单位,将域内由支持 TE 扩展的 IGP (OSPF-TE 或 ISIS-TE) 的信息收集上来,发送给控制器,由

控制器经过分析、整合后供流量调度、网络可视化等应用使用。

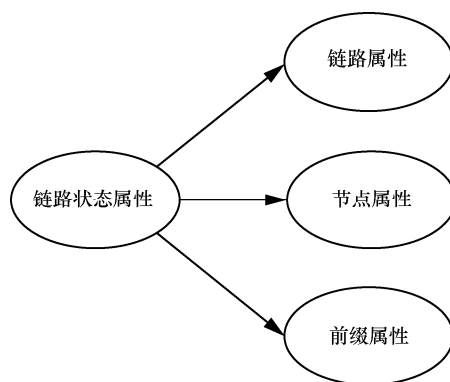


图 10 拓扑信息示意

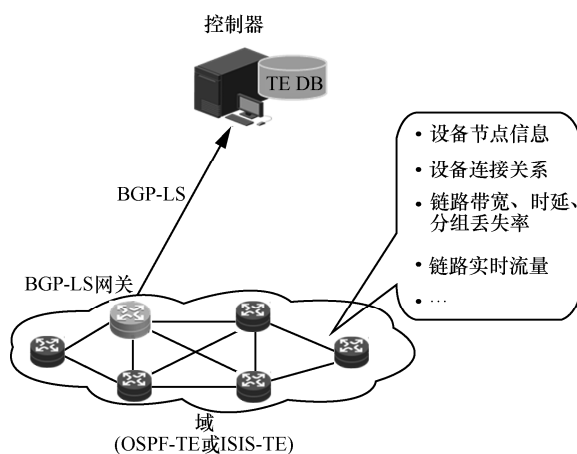


图 11 域内信息收集

5.2 高效的控制通道——BGP Flowspec

传统广域网中,管理员要实现复杂的流量控制,需要在设备上配置策略路由(PBR)来解决,但在广域网 SDN 中,控制器如果也通过下发 PBR 来进行流量控制,效率会很低,设备性能压力也很大,如果能采用路由协议来完成类似的功能,通过路由的发布和同步,就会解决性能和效率的问题,而传统的 BGP、MP-BGP 等协议只能基于目的地址进行报文控制,这时 BGP Flowspec 就成了一个比较合适的选择。

BGP Flowspec 是一种标准协议,在 RFC5575 中定义,它通过在 BGP 中定义新的 NLRI 允许携带源地址、端口号等类似于 ACL 的规则,并通过扩展团体属性(extended community)来定义匹配



字段执行的操作，如重定向、限速等。BGP Flowspec 是一种比较热门的技术，最初的目的是为了提供一种粒度非常细的对付 DDoS 的方法，但它也能替代 PBR，让控制器给设备下发策略路由，不但能够解决 PBR 在 SD-WAN 架构中存在的效率问题，而且能够充分利用 BGP 成熟的机制，很好地兼容现网设备，将 SD-WAN 能力快速引入现网，从而加快广域网的重构。

5.3 标准的控制通道——PCEP

在网络智能控制中，控制器要基于网络拓扑、流量情况等全局信息，并根据用户定义的应用 SLA 需求和调度策略，为各类应用计算最优路径，下发给网络设备，从而控制报文转发。目前用到的最重要的同时也是最标准的路径控制协议就是 PCEP (path computation element communication protocol)，在 PCEP 的定义中，控制器通常作为 PCE (path computation element)，网络设备作为 PCC (path computation client)。

- PCE：一个基于 TEDB (TE topology database) 计算带约束路径的软件模块，也可静态手工指定。PCE 可以位于网络路由器节点上，也可以部署在网络外的一台服务器（通常就是 SDN 控制器）上。PCE 可以是无状态 (stateless PCE) 的，也可以是有状态 (stateful PCE) 的。
- PCC：网络设备上的一个软件模块，向 PCE 发送路径计算请求，并从 PCE 接收路径计算应答。
- PCEP：PCC-PCE、PCE-PCE 间通过 PCEP 进行通信，传递计算请求和计算结果。

在路径计算时，可以由 PCC 发起路径申请（被动模式），也可以由 PCE 主动下发路径（主动模式）。

在 PCEP 最初的定义中，主要是传递 LSP 路径信息，随着 SR 协议的发展，PCEP 也进行了相应的扩展，这样，当控制器计算完路径后，转换

成具体的 SR 标签栈，就可以通过 PCEP 下发给设备，指导设备按标签栈指定的路径进行转发。

5.4 短、平、快的控制通道——NETCONF

标准协议有助于厂商互通和规模应用，但是一个协议的标准化是一个漫长的过程，任何一个小的新功能和扩展都要经过反复的讨论，周期往往比较长。有很多时候，用户的需求又需要一个业务短时间内快速上线，这时 NETCONF 就成了很好的选择。

NETCONF/Yang 作为目前最流行的网络管理协议，在 SDN 出现之前已经广泛应用，可以说不是什么新技术，但是标准协议，NETCONF 可以由设备厂商快速扩充新的特性，控制器只需要进行相应的适配开发即可，不用等协议标准化，大大地缩短了业务的上线周期，满足快速变化的应用需求，因此，它作为一种协议通道，在广域网重构和 SD-WAN 演进中，能够以短、平、快的方式，实现设备的自动配置和业务的灵活控制，起到举足轻重的作用。

NETCONF 是 IETF 于 2003 年 5 月为了解决 SNMP 在网络管理方面的问题而提出的，目前已有接近 20 个相关 RFC 标准。

NETCONF 是一种基于会话、面向连接的协议。NETCONF 协议在概念上主要分为 4 层，包括内容层、操作层、消息层、传输层，如图 12 所示。每个层分别对协议的某一个方面进行包装，并向上层提供相关的服务。分层结构让每层只关注协议的一个方面，实现更加简单，同时合理地解耦各层之间的依赖，将各层内部实现机制的变更对其他层的影响降到最低，各层功能介绍如下。

- 传输层：在客户端和服务端之间提供了一个安全通信路径。NETCONF 可以基于任何传输协议，但是对传输协议有一系列的基础要求，如要求必须面向连接、可认证，能确保数据一致性、私密性等。
- 消息层：提供了一种简单且与传输无关的成

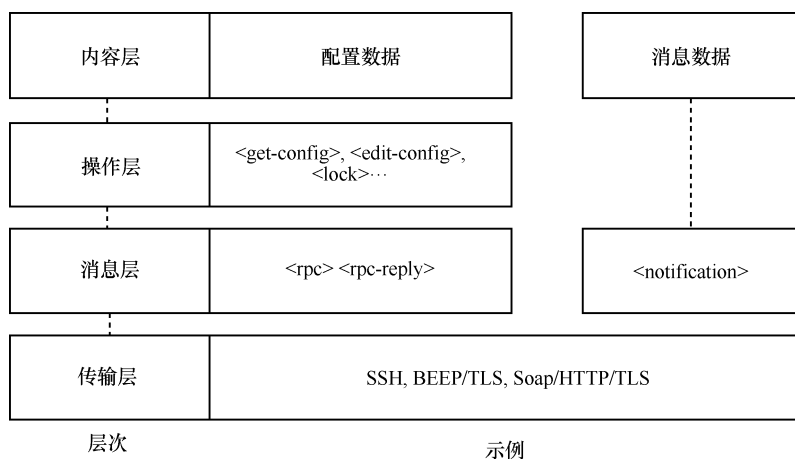


图 12 NETCONF 协议层次

帧机制。使用 RPC 相关消息对客户端和服务端请求和响应数据进行编码封装，通过 **notification** 消息对通知类数据进行编码。

操作层：定义了一系列基础协议操作集合，这些操作通过 RPC 方式进行调用，主要包括取值操作、配置操作、锁操作和会话操作等，其中 `get`、`get-config` 用来对设备进行取值操作，`edit-config`、`copy-config`、`delete-config` 用于配置设备参数，`lock` 和 `unlock` 则是在对设备进行操作时防止并发产生混乱的锁行为，`close-session` 和 `kill-session` 则是用于结束一个会话的操作。

内容层：表示被管理对象的详细内容。内容层的内容需要按照数据模型的方式进行定义，在 RFC6020 和 RFC6244 中定义了 NETCONF 的数据建模语言：**Yang**。

- 操作层：定义了一系列基础协议操作集合，这些操作通过 RPC 方式进行调用，主要包括取值操作、配置操作、锁操作和会话操作等，其中 `get`、`get-config` 用来对设备进行取值操作，`edit-config`、`copy-config`、`delete-config` 用于配置设备参数，`lock` 和 `unlock` 则是在对设备进行操作时防止并产生混乱的锁行为，`close-session` 和 `kill-session` 则是用于结束一个会话的操作。
- 内容层：表示被管理对象的详细内容。内容层的内容需要按照数据模型的方式进行定义，在 RFC6020 和 RFC6244 中定义了 NETCONF 的数据建模语言：**Yang**。

Yang 可对 NETCONF 数据模型、操作进行建模,覆盖了 NETCONF 协议中的操作层和内容层。Yang 因其为 NETCONF 量身定做,与其他的网络管理数据建模语言相比, Yang 在对 NETCONF 进行数据建模方面占有很大优势。

Yang 可对 NETCONF 数据模型、操作进行建模,覆盖了 NETCONF 协议中的操作层和内容层。Yang 因其为 NETCONF 量身定做,与其他的网络管理数据建模语言相比, Yang 在对 NETCONF 进行数据建模方面占有很大优势。

- **Yang** 可支持 NETCONF 所需要具备的所有特征，可完整描述客户端与服务器端交互的所有数据信息。
- **Yang** 在高层数据建模和底层数据编码间做

了良好平衡，使阅读者不仅可以清晰看到高层的数据模型，同时也很容易理解数据如何进行 NETCONF 操作编码，有很强的可读、可理解性。

- **Yang** 模块可无损转化为等价的 XML 格式，被称为 YIN（Yang independent notation），这使得相关应用可通过 XML 解析器或者 XSL 脚本进行操作。同时 YIN 也可以直接转化为 Yang，这为使用者带来了更多的便利性。
- **Yang** 是一种可扩展语言，允许标准制定者、设备商以及个人进行扩展定义，有很强的扩展性。
- **Yang** 并不期望解决所有问题，因此将问题解决重心限制在 NETCONF 数据模型上，这使得 Yang 虽然并不像 XML 的数据模型那样随意，但却更加简洁易用。
- **Yang** 与 SNMP 的 SMIV2 保持了兼容，基于 SMIV2 的 MIB 模块能够以只读的方式自动转换为 Yang 模块，提供了良好的兼容性。

- **Yang** 是一种可扩展语言,允许标准制定者、设备商以及个人进行扩展定义,有很强的扩展性。

- **Yang** 并不期望解决所有问题，因此将问题解决重心限制在 **NETCONF** 数据模型上，这使得 **Yang** 虽然并不像 **XML** 的数据模型那样随意，但却更加简洁易用。

- Yang 与 SNMP 的 SMIV2 保持了兼容，基于 SMIV2 的 MIB 模块能够以只读的方式自动转换为 Yang 模块，提供了良好的兼容性。

6 华三在广域网重构中的实践

6.1 新一代广域网方案——ADWAN

为了应对业界的发展趋势和满足用户需求，
华三率先开启了在广域网中的重构，为此，基于



SD-WAN 核心理念和开放架构推出了新一代广域网解决方案——ADWAN（application-driven WAN，应用驱动的广域网），整个方案以应用为中心，构建一个开放、智能、高效、易用的新一代广域网，承载日益丰富的应用流量，实现应用按需驱动网络、网络动态适应应用。

如前文所述，ADWAN 充分利用了当前重构广域网的关键技术，也可以说是这些关键技术的应用实践，整个方案采用 SD-WAN 思想，是一个分层、开放、灵活的网络架构，分为网络设备、控制器+App、管理编排 3 个层次，如图 13 所示。网络架构 3 个层次的详细介绍如下。

- 基础网络层：网络设备接收 SDN 控制器的控制和管理，除了传统的 SNMP、NETCONF、命令行等方式外，支持 SDN 架构下的 BGP-LS 协议、BGP Flowspec 协议、PCEP 等，和控制器进行通信；同时，在转发层面进行优化，支持 SR、VxLAN 硬件转发，提供高性能的转发平面。
- 控制器+App 层：整个方案基于开源的 ODL

平台，支撑各种 App 集成；根据广域网不同的场景，比如 DCI 网络、骨干网、分支接入等，开发定制化、场景化的 App，满足用户在不同场景下的网络需求；南向通过标准南向接口协议和设备互通；北向面向用户提供定制化的 API，实现和编排系统集成，满足用户差异化的业务需求。

- 管理编排层：通过调用 App 提供的 API，实现业务的策略定义和管理编排，全网的实时监控、可视化呈现及故障排查等，进而增强网络的可视化，简化网络的运维管理。

6.2 ADWAN 方案在 DCI 网络中的应用

随着业务云化的快速发展，数据中心越来越多，互访流量增加，并呈现出低时延、突发性和不均衡性等特点，承载这些流量的 DCI 网络变得越来越重要，同时面临的挑战也越来越多，如专线带宽利用率低，导致每 Mbit/s 成本很高；新业务不断产生的峰值流量，导致专线频繁扩容；传统流量工程复杂，很难全局端到端保障业务质量等，因此，急需对 DCI 网络的

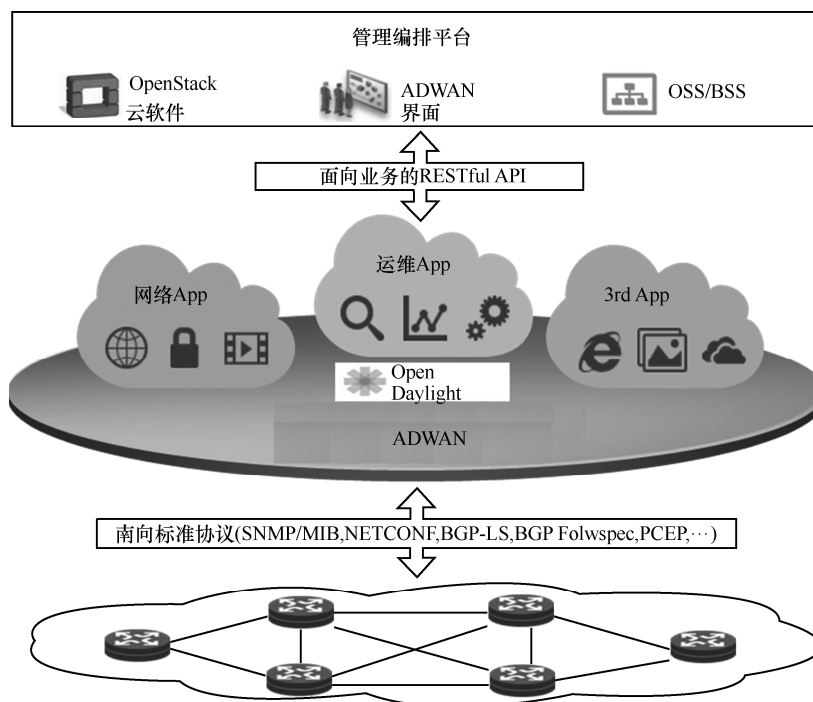


图 13 ADWAN 架构

流量进行集中控制、统一调度，ADWAN 方案很好地解决了这个问题。

ADWAN 在 DCI 中的架构如图 14 所示，首先将 DCI 网络由 MPLS 升级为 SR，使其同时具有 MPLS 和 SR 的转发能力，保留原有 MPLS 控制平面。其次，部署 SDN 控制器及 ADWAN App，作为 SR 的控制平面，指导设备的 SR 转发。最后，整个网络形成集中的 SR 和分布的 MPLS 两套控制平面，集中的 SR 控制平面作为叠加，优先级高，用于指导设备的报文转发；分布的 MPLS 控制平面作为备份，优先级低，用于 SR 控制平面失效后的逃生。整个方案既具有全局灵活的控制能力，又保证故障情况下的可靠逃生能力。

ADWAN 方案在 DCI 网络中为用户提供了以下功能。

(1) 网络可视化

通过 BGP-LS、SNMP/MIB、Netstream 等手段收集网络拓扑、链路状态及业务流量等信息，并对其进行分析和整合，然后对网络和应用进行可视化

呈现，形成 TE 数据库，为流量调度提供输入。

(2) 资源统一管理

根据收集到的网络拓扑，统一为设备和链路分配 SR 标签，通过 NETCONF 下发给设备。

(3) 业务配置自动化

通过 NETCONF 自动创建 SR-TE 隧道，将不同流量自动引到对应的隧道上进行承载，改变传统人工配置设备的方式，降低出错风险，缩短上线周期。

(4) 流量智能调度

通过对全网的专线资源池化和集中调度，根据不同应用的带宽和 SLA 需求，动态为应用选择最优路径，并定期对全网流量进行全局优化，均衡流量分布，提高专线利用率，降低线路成本。将控制器计算的路径信息转化为 SR 标签栈，并通过 NETCONF 下发给设备，控制设备报文转发。

当然，SDN 在广域网的应用才刚刚开始，ADWAN 在 DCI 的应用实践也是处于不断摸索和完善阶段，以下是笔者认为要考虑和改进的几点。

首先，DCI 网络承担着数据中心间的互访流

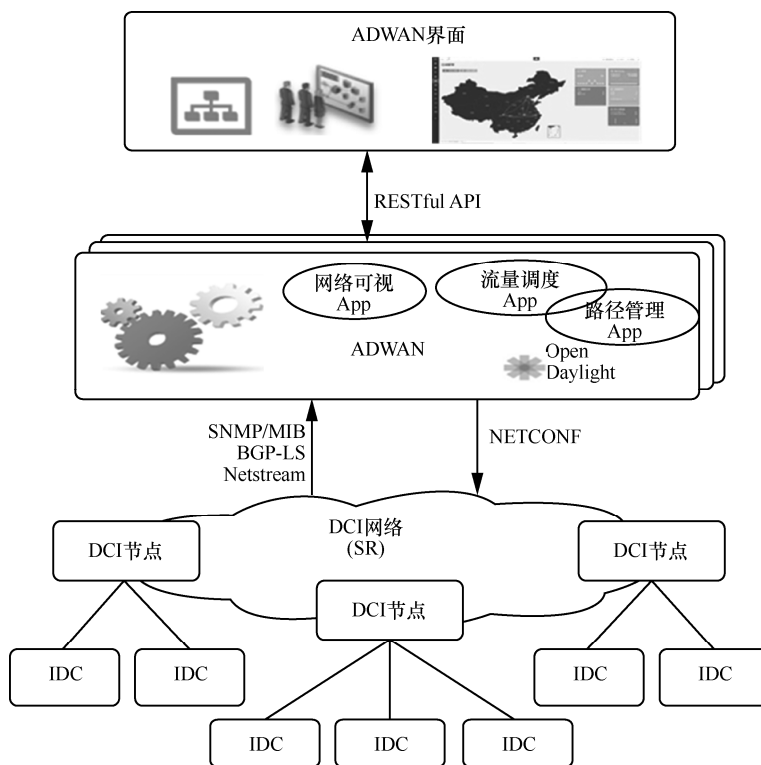


图 14 ADWAN 在 DCI 中的架构



量，它的稳定运行对用户的云计算业务至关重要，因此，可靠性是 ADWAN 方案在 DCI 网络中是否能够成功应用的关键，在设计和实施过程中，设计了从设备、路径、控制器、调度等多层次、全方位的可可靠机制，来保障出现各种故障后网络仍然能够正常运行，当然这些机制仅仅是开始，还需要在实际运行中进行检验。

其次，整个方案的性能也是是否能够成功应用的关键，包括控制器和设备的南向交互性能、控制器路径计算性能等，都是未来需要不断完善和提升的。

再次，流量调度依赖于控制器对网络运行状态、链路质量、业务流量等信息的准确性和及时性，但是目前采集和探测手段还是基于 SNMP、Netstream 等协议，不能很好地满足 SDN 环境下的要求，因此，未来需要采用新的协议，如 Telemetry，更准确、及时地获取网络信息，更好地为流量调度服务。

7 结束语

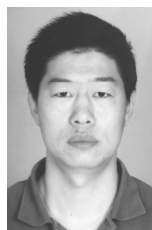
本文介绍了目前重构广域网使用的一些关键技术，既有核心的指导思想，也有关键的运行平台，还有重要的南向接口协议。广域网的重构，就是 SD-WAN 及相关技术在广域网中不断应用和实践的

过程，在这个过程中，既有对文中所述技术的改进，也有新技术的诞生，最终目标就是使广域网更加开放、智能、易用、可靠，同时又能够兼容传统网络，使用户的网络能够平滑迁移到 SD-WAN 架构上，快速满足用户不断变化的应用需求。

参考文献：

- [1] IETF. YANG-a data modeling language for the network configuration protocol (NETCONF): RFC6020[S/OL]. [2017-03-17]. <http://ftp.belnet.be/ftp.ietf.org/rfc/rfc6020.txt.pdf>.
- [2] Segment routing architecture[S/OL]. [2017-03-17]. <https://datatracker.ietf.org/doc/html/draft-ietf-spring-segment-routing>.

[作者简介]



郭晓军（1973-），男，新华三技术有限公司高级架构师，主要研究方向为广域网技术及解决方案。



万晓兰（1975-），女，新华三技术有限公司高级工程师，主要研究方向为网络及其新型技术。