

云网一体使能网络 即服务



Cloud Network Convergence Enables Network as a Service

朱海东/ZHU Haidong

(中兴通讯股份有限公司, 江苏 南京 210012)
(ZTE Corporation, Nanjing 210012, China)

摘要: 随着云业务的深入部署,除接入和连接性服务外,云网运营商还须需要将网络端到端带宽和时延等服务质量保障能力作为网络服务提供给最终客户。云网运营商基于云网一体网络架构,通过部署性能测量和故障发现、基于性能的路由、流量统计和状态预测,以及多域编排等系统,满足客户对不同需求层级能力的要求,使能网络即服务(NaaS)能力,实现网络与云化业务的有机结合。

关键词: 云网;多云;NaaS;服务等级

Abstract: In addition to access and connectivity services, cloud operators also provide multi-factor quality of service (QoS) assurance capabilities (such as end-to-end bandwidth and delay) as the network service to end customers. The cloud network service provider enables network as a service (NaaS) by building performance measure and fault detection, performance aware routing, traffic statistics and state prediction, and multi-domain orchestration systems to match the hierarchy of service needs upon the cloud network architecture.

Key words: cloud network; multi-cloud; NaaS; service-level agreement

DOI: 10.12142/ZTETJ.201902002
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.tn.20190409.1517.004.html>

网络出版日期: 2019-04-09
收稿日期: 2019-01-16

一直以来,网络规划人员采用自顶向下的设计方法构建并不断扩展数据通信网络,在过去的20年里成功地完成了业务和网络IP化过程。近年来,随着云计算的成熟应用以及信息技术(IT)行业与信息通信(IC)行业的交融,这种以需求为基础创建的网络通常无法如预期那样运行,也无法随网络规模的不断增长而扩展,无法满足客户的需求^[1]。未来网络的规划设计应以数据和数据中心(DC)为核心,围

绕云化业务的需求进行,这已经成为IT和IC行业的共识。云与网有机结合将构成云网一体的新型信息系统架构。

1 云网的概念

云网(CloudNet)将云计算架构与网络能力相结合。云网一体架构充分利用软件定义网络(SDN)和网络功能虚拟化(NFV)等技术提供的动态可编程能力,实现服务设施按需灵活部署。在该架构下,业务通

过跨Internet连接的若干数据中心协同提供,在提升资源利用率和业务可靠性的同时实现业务部署和管理的敏捷性^[2]。

1.1 云网一体参与方

在云网一体框架下,客户的业务由云业务服务商(CSP)和网络业务服务商(NSP)共同完成。CSP利用架设在数据中心之上的设备和资源,通过互联网或其他网络以随时获取、按需使用、随时扩展、协作共

享等方式,为用户提供数据存储、互联网应用开发环境、互联网应用部署和运营管理等服务。CSP通过云化数据中心提供服务,主要类型有软件即服务(SaaS)、平台即服务(PaaS)和基础设施即服务(IaaS)等。典型的CSP有阿里云、腾讯云、亚马逊AWS、微软Azure等。

除了公有云,互联网公司和政府、大型企业的私有云也是云的重要组成部分,如Facebook通过全球分布的数据中心为各区域客户提供服务,各级政府通过政务云提供便捷的公众服务。当私有云与公有云配合提供服务时,即成为混合云。

NSP则为客户提供网络接入服务、本地互连和跨地域骨干远程连接等服务,这些网络服务是综合电信运营商的基础电信业务之一。目前电信行业内一些运营商同时提供云服务和网络连接服务,转型为云网综合运营商,如中国三大运营商:中国电信(天翼云)、中国联通(沃云)和中国移动(移动云),它们都运营有各自的公用云和专有云业务;其他国家如西班牙电信(Telefónica)、德电(Deutsche Telekom)等传统电信运营商也纷纷进入公有云市场。

1.2 云网一体的研究范围

参照电气和电子工程师协会(IEEE)CloudNet委员会关注和研究的技术方向,云网一体架构分为以下几个技术领域^[2]。

(1)云网总体架构。

云网总体架构包括分布式数据中心架构、云数据中心大规模路由

组织方式、数据中心内/外部互访方式以及联合云和混合云组网等方面。业务的扁平化使得越来越多的数据中心网络引入Internet全球路由,实现跨地域的信息服务,这需要基于边界网关协议(BGP)路由属性和策略对流量流向进行优化控制。随着个人和政企业务上云的演进,联合云、混合云等多云方式部署也成为影响云网架构的重要推动力量。调研表明:81%的企业会采用多云策略,平均每个企业会使用5朵公用云和(或)私用云^[3]。大型互联网公司如Twitter也转向了混合云方案,将其冷存储和Hadoop业务迁移到了Google云上。在多云环境下,端到端业务需要在多个网络域和管理域间完成交互协同。

(2)云网资源管理。

云网资源管理根据业务需要动态按需提供计算资源、存储资源和网络带宽、服务等级等能力,涵盖的范围不仅包含数据中心基础设施和数据中心内部网络(DCN),还要包含数据中心互连网络(DCI),特别是通过多个NSP网络资源及CSP自用的骨干专网共同实现跨数据中心互联^[4]。云网资源管理需要获取物理网络和虚拟拓扑的全局视图和资源使用状态,提供最优的资源调度和使用策略,提供弹性、可扩展的动态资源管理方案。

(3)云网虚拟化技术。

云网虚拟化引入SDN转控分离架构和NFV虚拟网元,一方面可以实现业务部署的简化和自动化,提供敏捷服务,加速新业务上线速度;另一方面在多层、多域、多厂家

组网的复杂网络中提供端到端的管理和控制能力。分域化和层次化部署的SDN控制面还能够提供精细的控制粒度,提高系统资源利用率和运维效率。

(4)云网业务和云网安全。

云网业务和云网安全包括SaaS、PaaS、IaaS、大数据和数据分析、内容分发、多业务边缘计算等业务场景及其相关的安全要求。

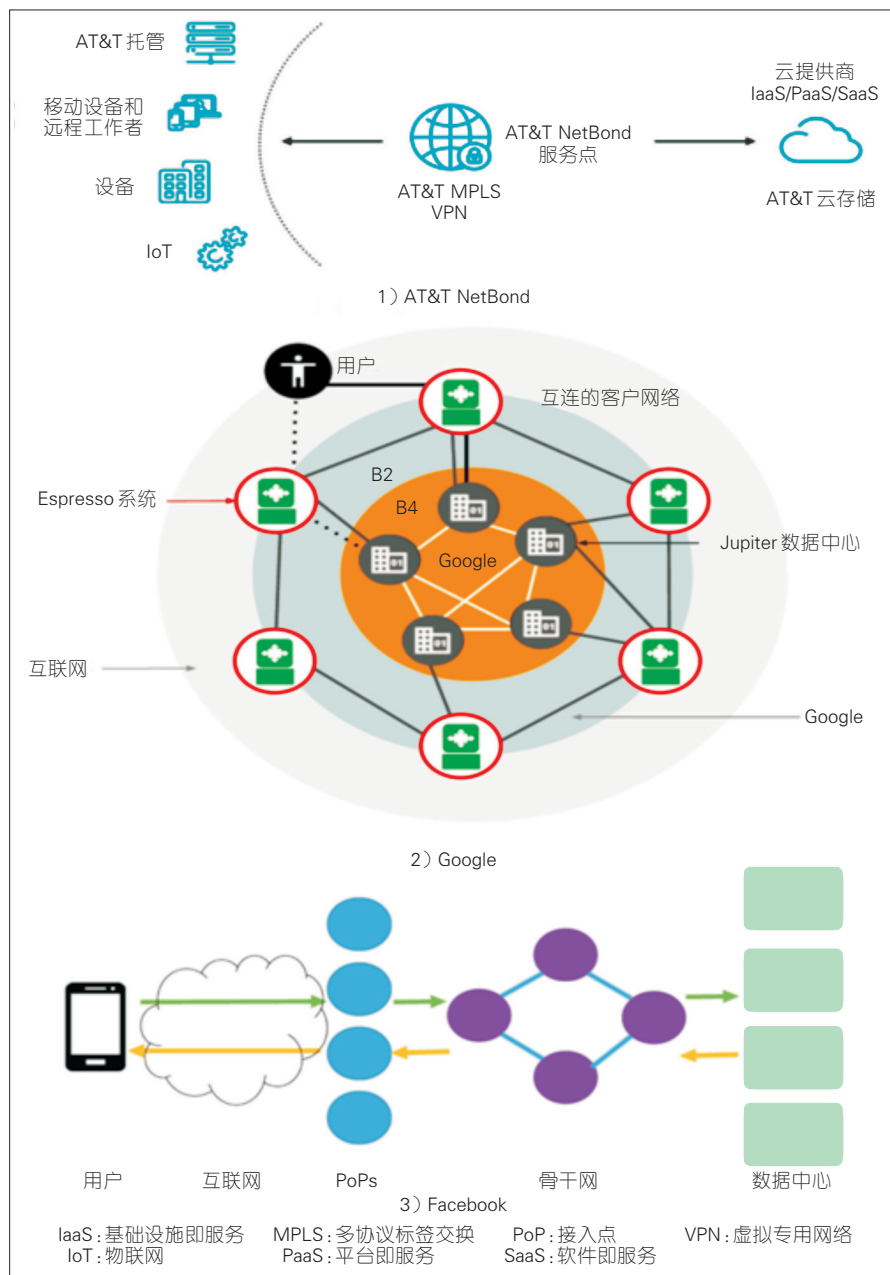
在云网一体架构下,为了保证端到端的业务体验,运营商还将提供给客户端到端带宽和时延等网络服务质量保障能力,从而实现网络即服务(NaaS)。

2 网络架构和业务需求

2.1 云网一体骨干网架构

对云网业务服务商而言,网络整体架构在很大程度上决定了业务组织方式和SDN/NFV等新技术的部署难度,从而影响资源优化利用所能达到的程度。从网络角度看,云网架构可以分为数据中心内外2个部分。数据中心内部DCN的目标架构主要有Fat Tree、Spine-Leaf等。数据中心之外的骨干网部分,得益于IP网与生俱来的互通性和可达性,目前NSP和CSP根据各自商业目标在其资源管控范围分别建设和发展各自的骨干网络系统。

如图1所示,AT&T针对企业入云提供了NetBond产品^[5],在网络中部署NetBond服务节点并与CSP背靠背接入,实现了AT&T企业多协议标签交换(MPLS)虚拟专用网络(VPN)和多云接入服务的整合,支



▲图1 AT&T、Google 和 Facebook 的网络

持用户接入私有云和公用云,还实现了带宽的按需扩展和连接的高安全保障。既是云服务商,又是内容服务商的 Google 拥有名为 B4 和 B2 的 2 张骨干网,前者由 33 个全球节点构成专用的大容量 DCI 网络^[6],后者在全球 70 多个区域完成 Google 数据中心与外部邻接运营商互通,

是终端用户访问 Google 的各种服务的必经通道^[7]。Facebook 的网络由边缘接入点 (PoP)、全球骨干和若干数据中心构成。用户通过 Internet 接入各地 PoP 点,再经由骨干网访问数据中心,Facebook 的骨干网同时承载外部用户流量和内部数据中心 (DC) 间流量^[8]。

从业务的视角看,上述 3 种骨干网架构没有把从最终用户到云中的、业务的全部端到端路径视为一个整体,实质上是在各自网络域内部进行的分段局部优化。根据 NSP 和 CSP 网络架构现状,用户端到端业务路径包含基础网络运营商 NSP 的接入/城域网、NSP-CSP 网间互通 PoP、运营商骨干网和 CSP 骨干/DCI 专网,以及数据中心内网 DCN 等。其中,互通 PoP 是新引入的网络节点,它不仅来完成传统 BGP 域间路由互通,还要满足业务互通对服务质量、安全隔离等要求,是实现业务端到端一致性的关键节点。当用户基于使用安全性和经济性等因素选择多网多云方案时,该方案还需要具备多个 NSP 和多个 CSP 间的多域资源组合优化和编排能力。图 2 分别展示了云网一体架构的单网单云和多网多云 2 种架构。

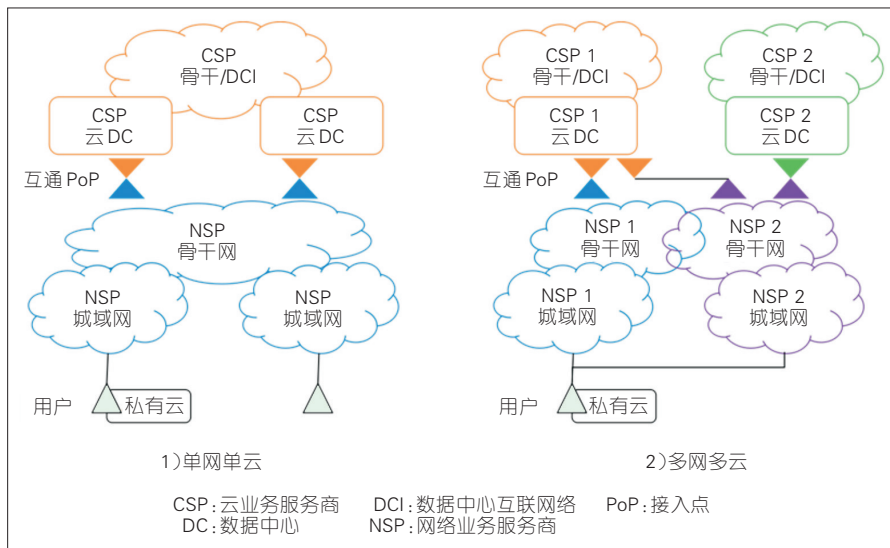
2.2 网络能力需求层级

业务运营对网络能力的需求存在若干层级,如图 3 所示。

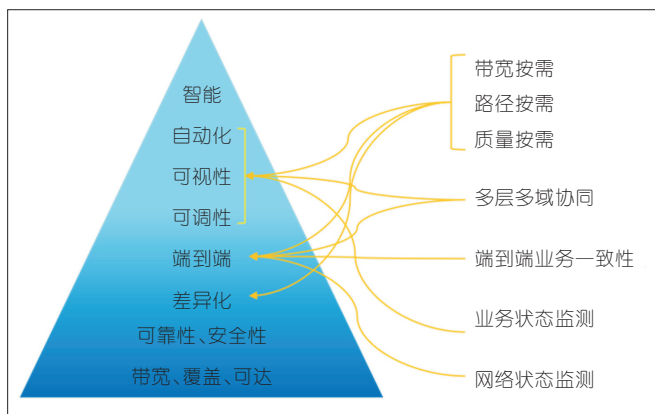
(1)最底层是网络物理带宽、物理覆盖范围和站点间可达性,这些属性是提供网络服务的基础。其中,物理带宽能力包含接入带宽和传输带宽能力。

(2)按照开放式系统互联 (OSI) 7 层模型,可靠性和安全性可分为物理层到网络层 (L1—L3)、传输层到应用层 (L4—L7) 2 部分。前者主要由网络能力保证,后者则由业务应用软件来实现。

(3)网络差异化业务保证能力是多业务综合承载的要求,即在一



▲图2 云网一体骨干网架构



▲图3 网络能力需求层级

张物理网络上按照不同业务的服务质量等级分配相应资源,不同业务对网络服务关键指标不同,比如企业跨DC大量数据复制业务关注在一定时间内总数据吞吐量,而对网络时延和丢包相对不很敏感。基于WEB的交互式应用,除了带宽和丢包率之外,影响传输控制协议(TCP)效率的时延因素成为业务保障的重要指标。目前,云网业务的主要技术指标有带宽、时延和可靠性等^[9],云网运营商在保证各等级业务服务达标的过程中,不仅为价值用户提供了更多产品选择,也使

得网络资源得以充分利用。

(4)网络的端到端能力是指从用户终端到应用服务器的全路径业务保障能力,特别是业务路径可能经过一个或多个运营商的多个管理域,在相邻的管理域之间需要保证用户体验的一致性。通过引入SDN编排系统,消除各个网络域采用的不同技术的形式差异,完成不同管理域间在资源定义、业务定义和服务等级协议(SLA)定义的映射和转换,同时实现资源的优化配置。

(5)自动化、可视性和可调性是云网敏捷能力的体现,也是实现

NaaS的灵活按需能力的基础。在SDN技术架构下,云网运营者将网络控制和转发分离,通过集中的控制面获取全局的资源 and 资源利用信息,并在路由层面实现对业务端到端全局路径组合优化。借助SDN网络规划和业务部署工具,云网运营者还可以实现业务部署的简化和自动化过程。只有具备了自动化、可视性和可调性的网络,才能作为一种服务发布给客户。

(6)智能化是更加高阶的自动化,其目标在于实现基于意图的网络(IBNS),从而为云网资源的设计、实施、运营提供覆盖全生命周期的自动化能力,将业务需求和网络资源配置实时同步。

3 关键技术和系统

3.1 性能测量和故障发现

云网业务的分布式天性使得最终用户的体验效果除了依赖应用软件外,还依赖于包括数据中心里的服务器、数据中心内外部网络以及运行在服务器上的虚拟化软件等组件的运行情况。从网络角度来看,云网运营商需要管理和保障数据中心内、数据中心间和用户到数据中心这3个网络域的业务带宽、时延和可靠性。研究表明当业务中断或体验下降时,用户很自然地将问题归结为终端问题和网络问题,而实际上50%的“网络”问题并不是由网络引发的^[10],因此,云网运营商需要有效的性能测量系统。

Ping探测是一种检测网络L3连通性和丢包率的通用手段。NSP

用 IP Ping 探测网络边界间的连通性, CSP 则使用 TCP Ping 和 HTTP Ping 检验业务层的连通性。显然后者更能符合用户端到端体验,但会存在 2 点问题:(1)针对网络中每用户每应用的测量不仅占用终端和服务端计算资源,同时其产生的海量测量数据需要存储和管理,例如 Microsoft Ping mesh 系统每天可以产生 2 000 亿个探针和 24 T 字节的数据^[10], Facebook 为了降低对资源的需求采用随机选择部分业务进行测量^[11];(2)将 CSP 应用层的测量结果应用到 NSP 网络层资源管理的映射关系是十分复杂的,特别是中间网络的路径上汇集了不同等级的海量应用,针对特定应用的网络问题根因分析缺少自动化手段。网络吞吐量和时延的性能测量分为带内和带外,带内的方式就是在报文头中设置或添加探针字段,在测量点统计性能指标,网络中间节点无感知。带外的方式则通过在网络中增加独立的探针数据流。由于数据网络的逐跳转发天性,这种方式无法保证探针流与业务流全程同路径。

基于业务的网络性能测量是支撑云网运维的重要手段,国际互联网工程任务组(IETF)正在研究、发展的 In-band 操作管理维护(OAM)采用带内机制,在数据包头增加 OAM 字段来实现对网络故障检测、路径测量、流量等的监测,有望成为各种隧道封装统一的测量机制。

3.2 基于性能的路由

基于路由的 L3 寻址是实现网络可达的基础手段。现有的内部网

关协议(IGP)和 BGP 协议在路由计算中对网络性能,特别是链路利用率、拥塞丢包性能、路径时延等实时动态性能变化并不敏感。云网业务通常跨越多个网络域,在域间运行的 BGP 协议为距离-矢量协议,其自身甚至对带宽也不敏感,所以依靠目前的 IGP 和 BGP 协议,无法针对业务的服务质量需求和网络性能的变化自动调整业务路由。为此,基于 IP/MPLS 和 SDN/Openflow 的流量工程隧道被引入到网络中完成流量流向的调整。隧道封装方式可以有很多种:虚拟局域网(VLAN)、虚拟可扩展局域网(VXLAN)、通用路由封装(GRE)、MPLS、分段路由(SR)等,云网运营商可以根据自身网络设备的支持情况和技术布局选择具体方式。

鉴于网络中业务种类的多样性和不同业务对网络资源需求的关联性,优化对象只能是网络中的高价值流量,或对整网带宽占用显著的、数量少流量高的“大象”流^[12]。针对这些高价值流量,应用上节讨论的性能测量结果,采用 SDN 集中控制方式调整流量路径,从而优先保障这些业务的质量。在同一个网络管理域中,可以为高价值流量建立端到端流量工程优化路径,当存在对丢包、时延等不敏感的背景流量时,网络带宽资源利用率可以接近 100%。从网络资源全局优化角度,业务等级数量越多不一定会使得优化效果更好。过多的业务等级使得资源分配算法更加复杂,流向调整引起的网络收敛速度也可能受到影响,甚至导致出现网络震荡。目前

在实践中,部署 2~4 个业务等级即可取得较好的调优效果。

3.3 流量统计和状态预测

基于性能的路由落地须要准确地统计当前在网业务流量并预测调整后的网络状态。使用简单网络管理协议(SNMP)可以以秒级粒度收集网络设备接口流量,业务级流量记录则需要使用 Netflow/IPFIX、sFlow 等基于采样的工具,云网的控制面需要将这些信息与网络拓扑结合,生成网络资源矩阵和流量矩阵。网络中某个路径性能与该路径上的负荷有关,而负荷是实时变化的。当负荷的变化主要是由控制面基于策略主动发起的,在一段时间内具有可预测性,使用流量工程 TE 技术进行网络资源优化才可行。这个时间段就是 TE 调整的窗口,在这个窗口期内,要根据网络性能完成 TE 策略的计算或下发生效。Facebook^[11]和 Google^[12]分别在边缘 PoP 和 DCI 场景实现了频度在 30 s 的调优方案。面对云网系统海量的业务流,Telemetry 技术提供效率更高的信息上报通道,可以提供更加精细的流量状态信息,进一步提高调优算法的效果。

3.4 多域编排

对一个由多域异构网络承载的业务,其网络转发面端到端路径可由多种制式的隧道拼接而成,控制面则归属于多个 NSP 和 CSP 的管理域范围。如图 4 所示,这需要引入多域编排(MDO)系统,完成多种技术和多个运营商间的资源管理、业

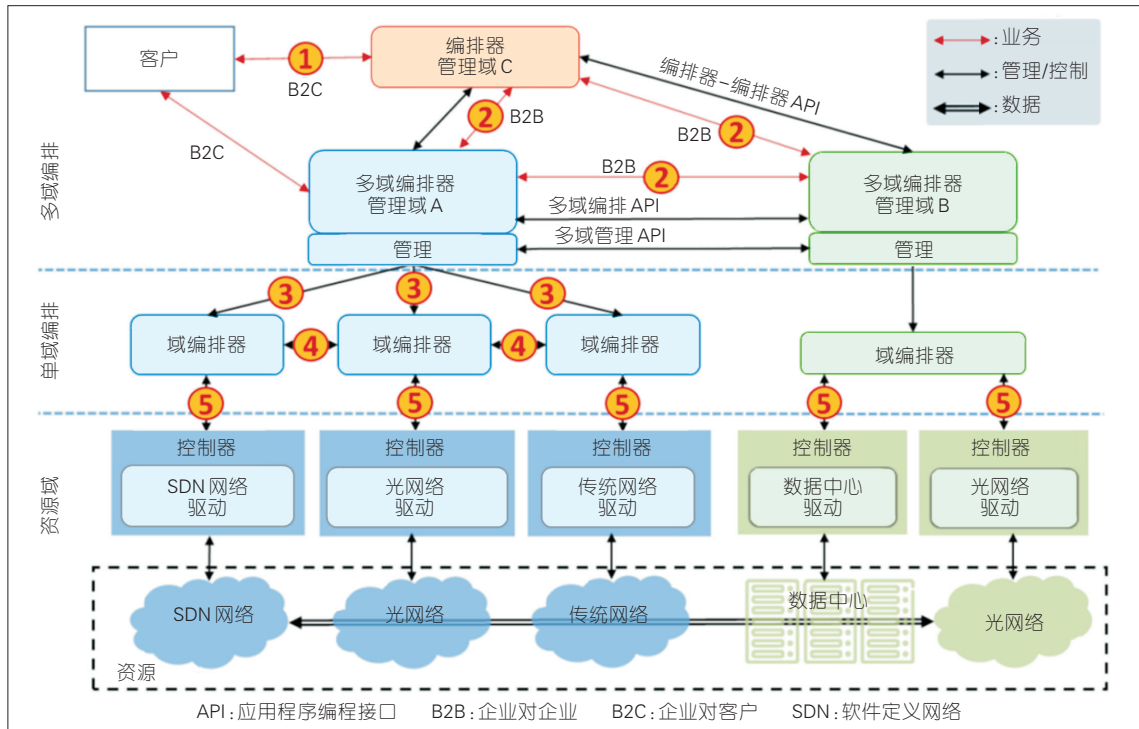


图4 多域编排系统框架

务管理和域间SLA协同^[13]。单域编排器从SDN控制器获取资源信息后进行本域的资源编排和业务编排,并与邻接域的单域编排器互通。多域编排器则通过应用程序编程接口(API)与本管理域中各下级单域编排器实现通信,完成多个子域间的协同编排,并实现对端到端业务的SLA保障。

4 结束语

在业务上云和多云时代,云网一体是ICT融合的必由之路。云网运营商发掘和扩展网络能力,将基础网络接入能力、连接保障能力、按需SLA能力与云化业务需求相结合,提供满足用户对高品质灵活服务的需要的新型网络服务,提升网络资源的商用价值,实现供需双赢。

参考文献

[1] OPPENHEIMER P. 自顶向下网络设计(第3版)

[M]. 北京:人民邮电出版社, 2011
 [2] IEEE- Cloudnet about CloudNet[EB/OL]. [2019-01-10]. <http://cloudnet2018.ieee-cloudnet.org/about.html>
 [3] 2018 State of the Cloud Report [EB/OL]. [2019-01-10]. <https://www.rightscale.com/lp/state-of-the-cloud>
 [4] NGUYEN C L, PING W, DUSIT N. Resource Management in Cloud Networking Using Economic Analysis and Pricing Models: A Survey [J]. IEEE Communications Surveys and Tutorials, 2017, 19(2): 954-1001. DOI: 10.1109/COMST.2017.2647981
 [5] AT&T NetBond Product Brief [EB/OL]. [2019-01-10]. <https://www.business.att.com/products/netbond.html>
 [6] CHIYAO H, SUBHASREE M, MOHAMMAD A. B4 and After: Managing Hierarchy, Partitioning, and Asymmetry for Availability and Scale in Google's Software-Defined WAN [C]//SIGCOMM 2018. Budapest, Hungary: ACM, 2018. DOI: 10.1145/3230543.3230545
 [7] KOK Y, MURTAZA M, JEREMY R. Taking the Edge off with Espresso: Scale, Reliability and Programmability for Global Internet Peering [C]// SIGCOMM '17 Proceedings of the Conference of the ACM Special Interest Group on Data Communication. Los Angeles, USA: 2017. DOI: 10.1145/3098822.3098854
 [8] SUNG E Y, TIE X, WONG H Y S, et al. Robotron: Top-down Network Management at Facebook Scale[EB/OL]. [2019-01-10]. <https://research.fb.com/publications/robotron-top-down-network-management-at-facebook-scale/>
 [9] CHRISTOPH A, AJAY B, EMILLIE D. Capacity Planning for the Google Backbone Network [EB/OL]. [2019-01-10]. <https://ai.google/research/pubs/pub45385>
 [10] CHUANXIONG G, LIHUA Y, DONG X.

Pingmesh: A Large-Scale System for Data Center Network Latency Measurement and Analysis[EB/OL]. [2015-08-23] [2019-01-10]. <https://www.microsoft.com/en-us/research/publication/pingmesh-large-scale-system-data-center-network-latency-measurement-analysis/>
 [11] BRANDON S, HYOJEONG K, TIMOTHY C. Engineering Egress with Edge Fabric: Steering Oceans of Content to the World [C]// SIGCOMM 2017. USA: ACM, 2017. DOI: 10.1145/3098822.3098853
 [12] ALOK K, SUSHANT J, UDAY N. BwE: Flexible, Hierarchical Bandwidth Allocation for WAN Distributed Computing[EB/OL]. [2019-01-10]. <https://ai.google/research/pubs/pub43838>
 [13] RICCARDO G, DAVID P, PAOLO M, et al. Multi-Domain Orchestration and Management of Software Defined Infrastructures: a Bottom-Up Approach[C]// 2016 European Conference on Networks and Communications. IEEE Communications Society. USA: IEEE, 2016

作者简介



朱海东, 中兴通讯股份有限公司产品规划总工; 主要从事数据通信产品和网络规划设计等工作。