

算网底座构建的探索与实践

Exploration and Practice on Computing Power Network Base Building


秦壮壮¹,赵金水¹,屠礼彪²,王 笙¹,刘 敏¹(1. 中国联通北京分公司,北京 100052;2. 中国联合网络通信集团有限公司,北京 100033)

Qin Zhuangzhuang¹,Zhao Jinshui¹,Tu Libiao²,Wang Sheng¹,Liu Min¹(1. China Unicom Beijing Branch, Beijing 100052, China; 2. China United Network Communications Group Co., Ltd., Beijing 100033, China)

摘要:

构建算网底座是实现云网协同、算网融合的下一代算力网络架构的基础。算力网络具有泛在接入、超大联接、灵活可编程、确定性、感知业务和安全可信等需求。对算网底座的架构进行了研究,提出了骨干、汇接和接入的分层架构设想及云枢纽局、云汇接局和云端局3类基础设施。介绍了某市联通基于“IPv6+”技术重构传统IP城域网,构建资源层、管控层和服务层3层云汇接算网底座的实践及组播承载、算力优先网络等应用成果。

关键词:

算力网络;云网协同;IPv6;算力优先网络
doi:10.12045/j.issn.1007-3043.2022.04.004
文章编号:1007-3043(2022)04-0016-07
中图分类号:TN919
文献标识码:A
开放科学(资源服务)标识码(OSID): 

Abstract:

The construction of a computing-network base is the foundation for realizing a next-generation computing power network architecture with cloud-network collaboration and computing-networking convergence. Computing power networks have the requirements of ubiquitous, super connection, flexible programmability, determinism, service awareness and security and trust. The architecture of the computing network base is studied and a layered architecture of backbone, aggregation and access and three types of infrastructure: cloud junction bureau, cloud exchange bureau and cloud end bureau, are proposed. Then it introduces the practice of reconstructing traditional IP MAN based on "IPv6+" technology in a city branch of china unicom, building a three-layer cloud exchange computing-network base with resource layer, control layer and service layer, and application results such as multicast carrier and computing first network.

Keywords:

Computing power network; Cloud-network collaboration; IPv6; Computing first network

引用格式:秦壮壮,赵金水,屠礼彪,等.算网底座构建的探索与实践[J].邮电设计技术,2022(4):16-22.

1 概述

ITU 成立的 Network2030 焦点组(FG-NET-2030)对网络 2030 的定位是“从连接人,到连接组织,到连接社会”。随着人工智能(AI)的出现和处理器处理能力的大幅提升,机器算力将作为重要的生产力推动人类社会进入“智能社会”。“智能社会”遍布世界的传感器会产生海量的数据,对网络提出泛在、超大连接的传送需求。例如脑机接口预计能同时读取 100 万神经元

信息,未压缩数据传输速率可达 Tbit/s。

国家发展改革委会同有关部门研究制定了《全国一体化大数据中心协同创新体系算力枢纽实施方案》,提出布局建设全国一体化算力网络国家枢纽节点,启动“东数西算”工程,构建国家算力网络体系。

2021 年 12 月,国务院印发《“十四五”数字经济发展规划》,在“优化升级数字基础设施”方面,明确要求推进云网协同和算网融合发展。加快构建算力、算法、数据、应用资源协同的全国一体化大数据中心体系。推动智能计算中心有序发展,打造智能算力、通用算法和开发平台一体化的新型智能基础设施。

收稿日期:2022-03-06

2022年2月,国家发改委等四部门复函同意粤港澳大湾区、成渝地区、长三角地区、京津冀地区启动建设全国一体化算力网络国家枢纽节点。全国一体化大数据建设将成为中国数字经济发展的新基座,国家算力网络将成为数据要素流通、畅通经济循环的新通道,国家枢纽节点将成为培育经济增长新动能、支撑国家大数据战略的新支点。

面对挑战与发展机遇,中国联通提出了“数字信息基础设施运营服务国家队、网络强国数字中国智慧社会建设主力军、数字技术融合创新排头兵”的新定位,定义“大连接、大计算、大数据、大应用、大安全”五大战略业务,挺进数字经济主航道、构建数字经济连接底座基础布局。

2 算力网络的需求

算力网络需要满足泛在接入的需求。算力网络和5G、Wi-Fi 6、F5G等新一代接入技术相结合为海量传感器终端提供接入服务。

算力网络需要满足海量数据传送的需求。未来10年接入数据速率将从千兆向百吉比特超连接演进,IP承载网和光传送网要支撑百倍的容量提升。

算力网络需要满足灵活可编程的需求。大型算力基础设施布局相对集中,温冷数据“东数西算”,边缘算力将分担热数据的算力需求。因此算力网络要连接云化集中和边缘分布的多级算力基础设施,为数据在算力资源池之间的调度提供灵活敏捷的可编程路径。

算力网络需要满足确定性服务的需求。分布式计算和存储对通信网络提出了确定性时延、高可靠性和稳定性等要求。传统IP承载网尽力而为的转发方式缺乏严格有效的服务水平协议(service level agreements, SLA)保证手段。

算力网络需要满足云网协同、算网融合、网络服务化的需求。从效率和能耗的角度,专业化、弹性的算力资源池将是未来数字化社会的新型关键信息基础设施。网络管道的承载能力要以服务的形式向算力开放,实现算网一体融合调度。

算力网络需要具备感知业务、安全可信的能力。网络成为数据和算力之间的供需平台,为算力和数据的供需提供流通路径和可信服务,算力网络向上感知算力使用情况,向中感知业务,向下感知网络资源,为数据提供差异化输送管道,可进一步提升算力使用效

能,实现“一键网调云”。

3 算网基座的架构设想

算网中的算力供给侧有很强的分布式计算属性。“东数西算”必然产生全国或区域性的省际数据调度场景。边缘计算作为低时延解决方案和算力枢纽的有效补充会产生本地网间和本地网内的流量调度场景。

由于调度算力的成本远高于调度数据,因此算网中的算力消费侧需要网络具备根据时延、带宽、负载、能耗和SLA等约束条件灵活选择算力资源的能力。

根据对中国联通新一代数字基础设施CUBE-Net3.0网络创新体系的研究和学习,以及对CUBE-Net3.0算力网络架构的理解,某省联通提出如图1所示的简洁、智能、开放的分层算网基座架构设想。

3.1 算网基座的网络分层模型

从网络层级角度,算网基座自上至下分为云骨干、云汇接和云接入3层。

a) 云骨干负责疏导云汇接本地网之间的流量,接入国家算力枢纽、区域性算力中心和超级算力。云骨干可进一步分为云国家骨干(国干/一干)和云省骨干(省干/二干)。现阶段云省干存在的主要目的是从骨干网卸载云间流量,简化跨域开通难度,便于省级网络SDN控制器与省云管编排层协同,缩小管理域,提升开通效率。今后随着网络规划趋于简洁和扁平化,以及自研控制器能力的提升,也可参考169网取消云省干层,各云汇接本地网直接接入云国干。

b) 云汇接设置于本地网内,负责为城域内的云间、入云流量提供疏通管道,并作为访问本地网以外算力资源的关口。云汇接应尽量保持技术先进和架构简洁,省云池、通信云、区域和布局型边缘算力、行业云、客户私有云等算力资源可以直接接入云汇接网络。建议允许省分公司根据自身资源情况和网络发展水平选择构建云汇接的范围和方式。

c) 云接入负责接入算力消费者和公有云资源,并把入云数据转发给云汇接。云接入的目标是尽量提供更大的算网基座覆盖范围和更快的业务拓展速度,不必拘泥于网络层级和采用何种技术架构,建议将传统IP城域网、智能城域网、IPRAN和各类刚性管道都纳入云接入的范畴。

3.2 算网基座基础设施分类

从基础设施角度,算网基座自上至下分为云枢纽

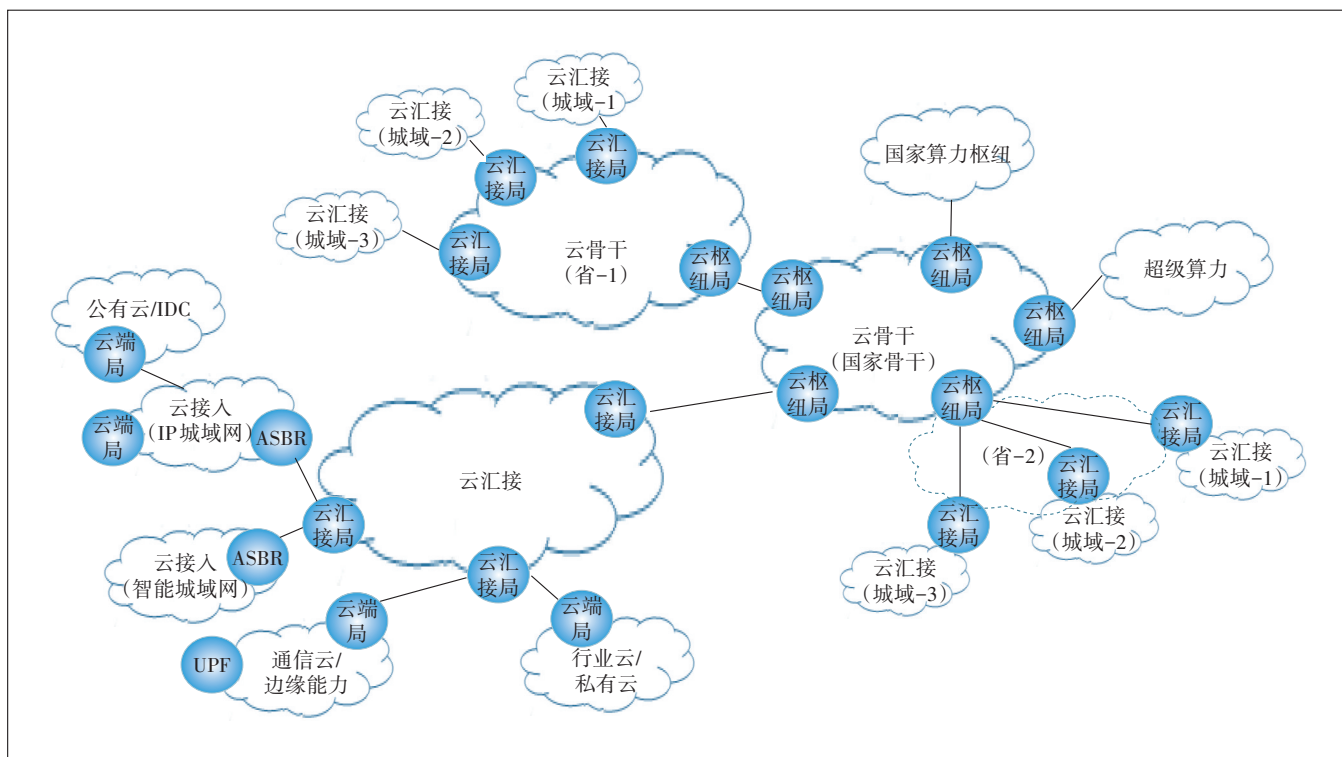


图1 分层算网基座架构

局、云汇接局和云端局3类。

a) 云枢纽局是云骨干的算网基础设施。云枢纽局内可以设置云骨干、169、中国联通产业互联网(China Unicom Industry Internet, CUII)路由器、传送网干线ROADM等骨干算力网络设备和骨干云池等骨干算力资源。

b) 云汇接局是云汇接的算网基础设施。云汇接局可以设置云汇接网的云运营商(provider, P)路由器、云运营商边缘(provider edge, PE)路由器、随选路由器和自治系统边界路由器(autonomous system boundary router, ASBR)等网络设备。

c) 云端局是云汇接网络的延伸。需要接入云汇接的算力消费者或算力资源可以采用直接接入云汇接局或通过云端局接入云汇接网络的方式。由于基于IPv6转发的段路由(segment routing over IPv6, SRv6)技术的灵活性和可扩展性,云端局也可以设置在云接入网络内,云端局和云汇接局之间通过SRv6隧道进行通信。

4 某省联通构建算网基座的实践

4.1 总体目标

通过验证和部署以SRv6、下一代组播、随流检测

和感知应用的IPv6网络(application-aware IPv6 networking, APN6)等为代表的“IPv6+”技术,以及创新灵活可编程、安全可信、时延可确定和应用可感知的“四可”新型城域网,构建满足数字化转型所需的云网协同、算网融合的智能算网基座。

4.2 建设思路

某省联通城域网是超大规模的扁平化单一本地网,没有建设云骨干的必要性。

综合考虑业务需求、资源情况和现有IP城域网、IP承载网的定位与发展水平,选择采用“IPv6+”技术重构传统IP城域网成为云汇接网络,新建云汇接管控系统的方案。这种建设思路的优势在于:

a) 选择较高的技术起点。某省联通IP城域网正处于扁平化和国产化转型期,IP城域网老旧设备存量、厂家众多、平台能力参差不齐,SDN化改造程度低。采用“IPv6+”技术重构可以将云汇接直接跃升到部分具备自动驾驶、云网协同能力的目标网,也有利于下一步向IPv6单栈演进。

b) 提供更快的交付速度。某省联通5G承载网选择先从IPRAN升级到智能城域网设备,再从MPLS升级到SRv6的技术发展路线,较好较快地满足了5G发展对承载网的能力诉求,但同时也给SRv6改造工作带

来了一定的难度。在IP城域网基础上重构云汇接网络不仅可以快速形成供给能力,尽量减小对智能城域网改造的干扰,还可以确保现阶段5G业务的优先承载。

c) 充分发挥在“IPv6+”创新领域的经验积累。某省联通是IPv6改造部署的坚定践行者和“IPv6+”技术创新的排头兵,近年来在若干领域第一个应用“IPv6+”技术,建立了系统内首个“IPv6+”联合创新实验室,国内首批通过“IPv6+”Ready认证的单位。正是这些经验积累让某省联通具备了建设高水平算网基座的信心和底气。

d) 便于快速迭代和敏捷创新。保持云汇接的紧凑规模和简洁架构,可以充分利用IPv6高速发展的红利,快速迭代采用新技术、落地新场景。云汇接目前已成为某省联通云网协同应用和产业互联网创新应用的试验床。

e) 有利于新技术的复制和推广。云汇接网络是IP城域网的一部分。云汇接采用的各项先进技术可以快速在IP城域网内复制和推广,带动IP城域网整体演进升级。

4.3 某省联通云汇接整体架构

某省联通云汇接参考CUBE-Net3.0网络创新体系架构,自下而上分为资源层、管控层和服务层,如图2所示。

资源层在扁平化IP城域网架构基础上参考智能城域网模型,其目标是构建以算力资源为中心的云汇接网络。云汇接局内设有云P、云PE和云随选设备,随业务互通需求可设置ASBR。在5GC通信云节点、MEC边缘云节点、IT云/行业云节点,设置专门的云POP(point of presence)设备接入云汇接网络。云汇接网络是为云间、入云和云组网业务提供服务的网络基础设施和底座。

管控层由各种网络控制器组成。目前部署1套云汇接控制器和1套超级控制器。网络控制器负责从网络设备内置的传感器收集状态、计算网络路径、下发业务指令、提供人工智能辅助的开发运维服务等。管控层南向对接资源层,北向为服务层提供能力开放接口。管控层在东西向上还可以提供跨专业、跨厂家的网络协同。

服务层完成云网协同业务的编排以及云网资源的协同和分配。服务层将云网能力以服务的形式统一提供给业务使用,对产品创新和交付人员屏蔽了实

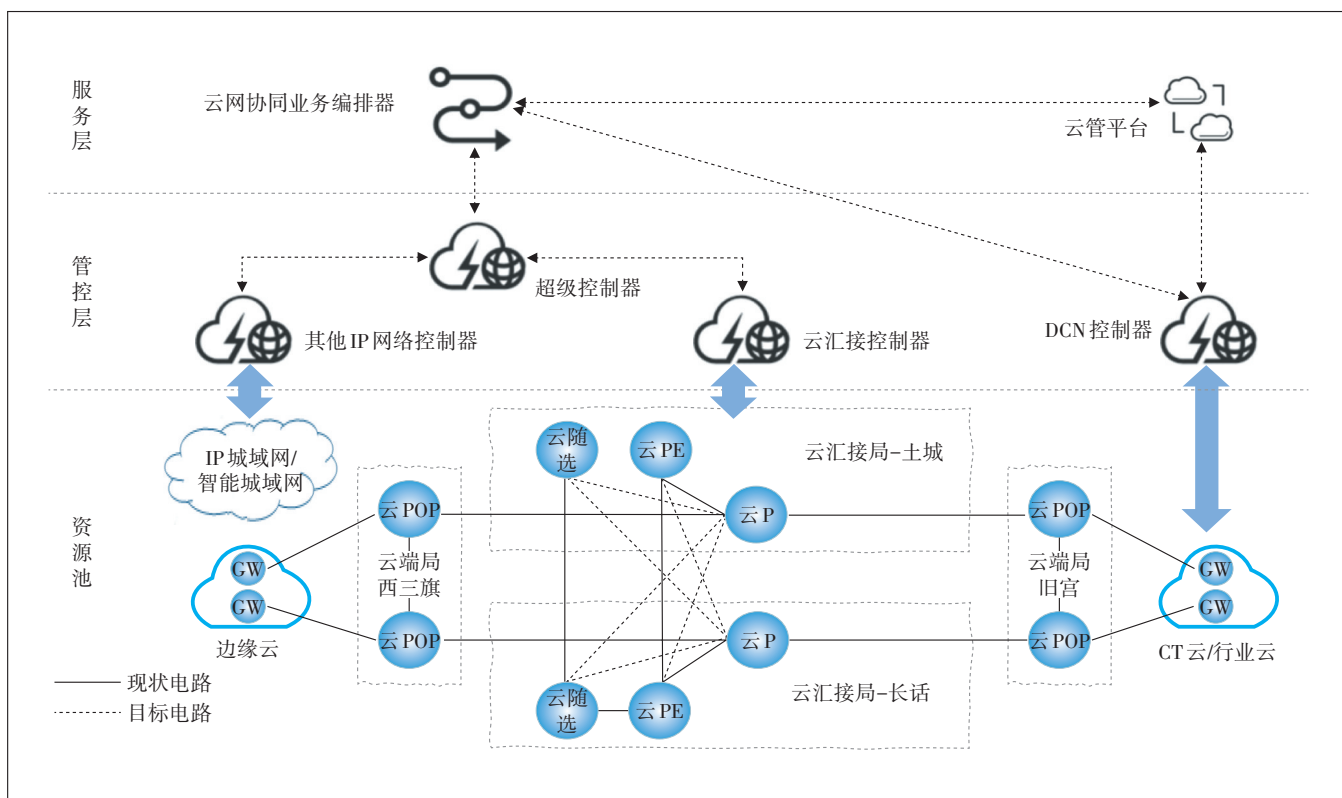


图2 某省联通云汇接整体架构

现细节, 明显提升业务发放的效率。

5 算网基座的关键技术

某省联通算网基座全面采用“IPv6+”系列技术, 是面向云网协同、算网融合的下一代互联网实践。

5.1 SRv6 源路由技术实现网络可编程

SRv6 是基于源路由理念而设计的在网络上转发 IPv6 数据报文的一种协议。SRv6 通过在 IPv6 报文头中插入一个路由扩展头 SRH(segment routing header), 在 SRH 中压入一个显式的 IPv6 地址栈, 并由中间节点不断地执行更新目的地址和偏移地址栈的操作来完成逐跳转发。目前 SRv6 的网络编程架构和分段路由扩展头都已经发布为国际互联网工程任务组(IETF) 的征求意见稿(RFC)文档。

5.2 随流检测技术实现业务质量可感知

随流信息遥测(in-situ flow information telemetry, iFIT) 检测技术基于 RFC 8321 (alternate-marking method for passive and hybrid performance monitoring) 实现, 是一种对实际业务流进行特征标记(染色), 并根据特征字段进行丢包和时延监测的技术。

iFIT 技术可以随流逐包精准检测业务的时延、丢包、抖动等性能信息, 并通过秒级数据采样实时呈现。采用逐跳部署模式, 可以还原业务路径, 快速进行故障定界定位。

5.3 下一代组播技术实现随享云播

比特索引显示复制(bit index explicit replication, BIER) 技术采用节点比特标识来复制组播报文。中间节点根据组播报文中携带的比特串判定如何复制和转发, 无需保存基于流的组播状态。

基于 IPv6 的比特索引显示复制(bit index explicit replication over IPv6, BIERv6) 进一步采用 IPv6 替代

MPLS 作为数据转发面, 利用 IPv6 扩展报文头的天然可编程能力携带组播业务 VPN 信息和 BIER 转发指令, 彻底摆脱了要单独分配 MPLS 标签的转发机制。

BIERv6 结合 SRv6, 可以实现在 IPv6 单一数据转发面上同时承载单播和组播 VPN, 真正做到全业务综合承载。

6 算网基座的应用

某省联通已建成云汇接算网基座, 实现了云网业务的一点发放、同开同停, 具备了云边协同、多云汇接、一线入多云、一线多业务、随享组播等多个业务场景能力, 支撑政企商业单元(business unit, BU) 推出行业云网、融汇云边、智慧企务、随享云播、All-In-One 等 5 个产品, 提升了该省联通政企产品的市场前景性和竞争力。

云汇接还为“IPv6+”技术的验证和算网产品创新思路的落地提供了基础条件。

6.1 云汇接承载智慧安防业务

某省联通采用云汇接网络承载部分智慧安防业务, 可以实现边缘云和中心云存储之间的资源协同, 对内外提供统一的智慧安防平台服务(见图 3)。

安防视频处理服务端、监控集成平台和存储服务集中部署在云端并接入云 POP。摄像头等终端通过端局交换机接入云 PE。边缘服务端就近部署, 忙时本地存储, 闲时上传集中云存储。门禁鉴权和视频流等需要隔离和时延敏感的业务可分配到边缘算力就近进行处理。

6.2 云汇接承载随享云播业务

某省联通依托于云汇接和 BIERv6 技术推出了随享云播业务。随享云播能够为客户提供面向 8K 的下一代组播虚拟专网服务, 如图 4 所示。

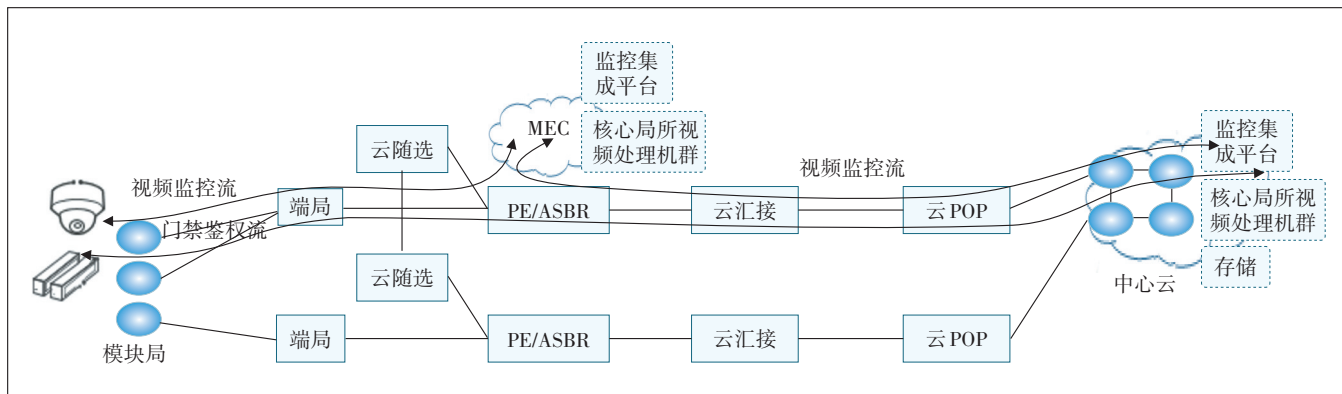


图 3 基于云汇接的智慧安防平台方案

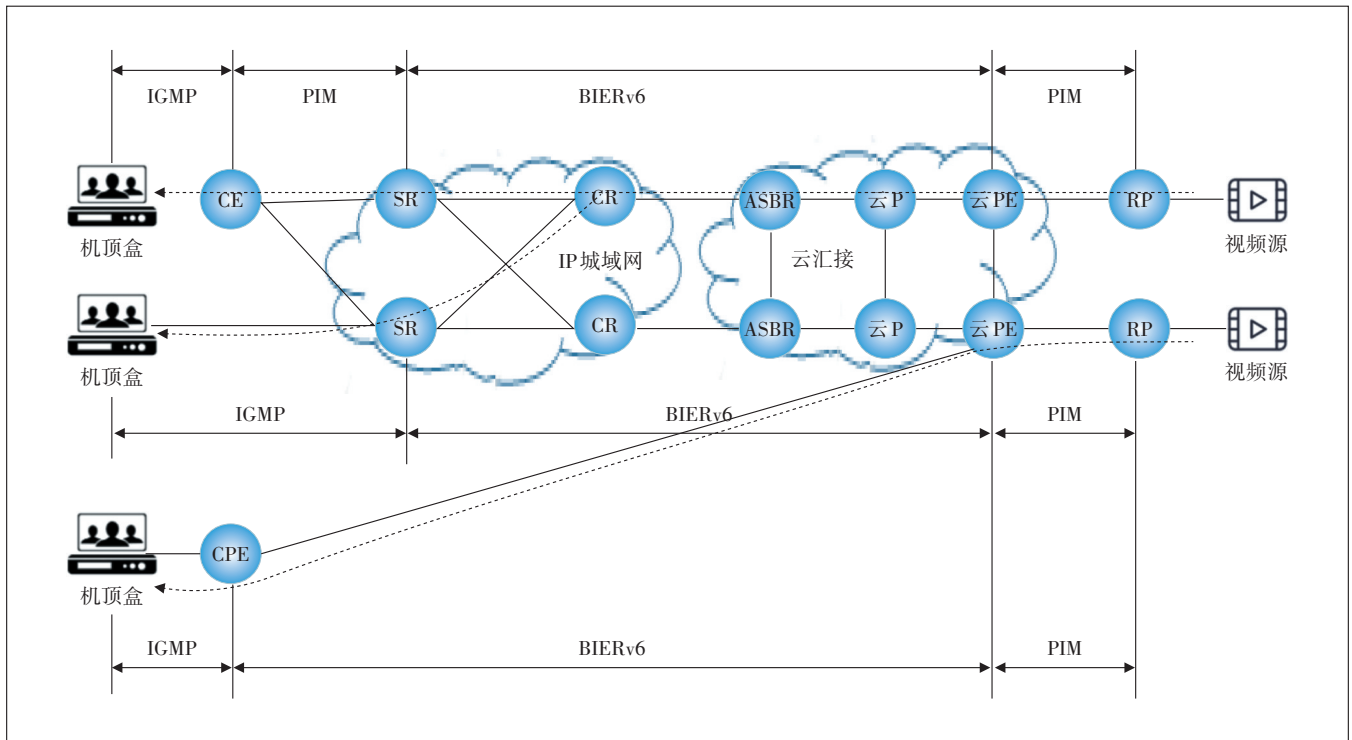


图4 基于BIERv6的随享云播

客户汇聚点(Rendezvous Point, RP)接入该省联通的云汇接网络,接收端接入该省联通任何一张IPv6可达的网络。RP设备引组播流进入云汇接网络,云PE作为根节点,客户终端设备(customer premise equipment, CPE)或者业务路由器(service router, SR)设备作为叶子节点,采用BIERv6为客户的媒体流提供组播虚拟专网服务。当前组播复制点设置在云汇接内,随着IP城域网设备的升级,可将组播复制点下沉到城域网核心路由器(core router, CR)或SR。

6.3 算力优先网络的验证

某省联通云汇接网络初步构建了算网基座的网络底座,为算网融合的技术验证提供了试验床(见图5)。算力优先网络的验证(computing first networking, CFN)是目前正在研究的一种算网融合控制面实现方式。CFN能够感知算力资源的使用情况并通过边界网关协议(border gateway protocol, BGP)进行传递,可以识别算力属性的云汇接网络设备就能够将数据引导到相应的算力资源池。与“云调网”的SDN集中控制方式不同,CFN是一种分布式算力路由技术,符合“网调云”的思想。

某省联通配合某研究院验证了CFN在智能安防场景下的应用。在该场景下,安防摄像头捕捉到的图

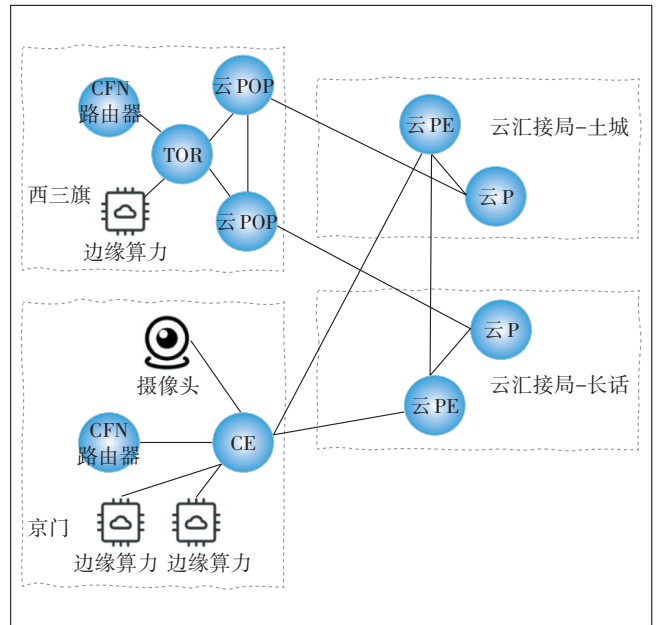


图5 CFN云汇接验证环境

像被发送到2个MEC边缘算力上进行AI识别。试验证明采用CFN可以由网络主动根据2个算力站点的负载情况,将图像数据调度至低负载站点处理,而且整个过程对摄像头端和图形处理器(graphics processing unit, GPU)算力端都是透明无感知的。

7 存在的问题与解决建议

某省联通在构建算网底座的过程中,详细分析了遇到的问题,并提出了解决建议。

算力网络的概念提出以后,业界对算力的具体供给形式、算力对网络的量化诉求、算网融合和云网协同之间的关系等关键问题尚未有统一认识。某省联通选择从资源层入手,以云网协同为抓手构建管控层和服务层,努力建设灵活可编程、感知可确定、安全可靠、应用可感知的智能云汇接算网底座。

云汇接算网底座建成投产后已成为某省联通 IPv6 新技术验证、云网协同产品创新和数字化转型的重要新型信息基础设施。随着业务需求的增加,保持网络简单敏捷和提高网络覆盖规模成为必须要解决的矛盾。该省联通将对云汇接的投资建设采取审慎态度,根据自身资源情况设置合理的算网底座发展目标。

确定性正逐渐成为 IP 网发展的痛点。某省联通在智能城域网中提供基于 FlexE 和 SRv6 + EVPN 技术的承载网切片。FlexE 成本高、互通性差,互联网业务流量大、流向复杂,这些因素限制了 FlexE 在传统 IP 城域网中的应用。某省联通将积极跟进确定性 IP 技术的发展,在适当的环境下进行验证。

8 结束语

云汇接算网底座以可编程网络做深大连接,进而为某省联通下一步做强大计算、整合网络与算力资源,推进架构先进、安全可靠、服务卓越的算力网络新布局提供有力的支撑。某省联通 2022 年云汇接的工作重点是面向产品业务需要,优化云网协同对接能力,满足一线入多云、云边协同、随享云播等重点场景的网络能力要求。在迭代创新方面,要逐渐补全算网底座所需的功能要素,充分发挥“IPv6+”联合创新实验室的平台和融通作用,产学研用相结合推动 IPv6 网络高质量发展。

参考文献:

[1] 秦壮壮,屠礼彪,臧寅,等.基于“IPv6+”的5G承载网切片技术与应用[J].电信科学,2020,36(8):28-35.
[2] 曹畅,唐雄燕,张帅,等.算力网络:云网融合2.0时代的网络架构与关键技术[M].北京:电子工业出版社,2021.
[3] 中国联合网络通信有限公司研究院,中国联合网络通信有限公司广东省分公司,华为技术有限公司.中国联通CUBE-Net 3.0网络创新体系白皮书[R/OL].[2022-01-04].https://www.doc88.com/

p-03147126445834.html?r=1.
[4] 中国联通研究院.算力网络架构与技术体系白皮书[R/OL].[2022-01-04].https://xw.qq.com/cmsid/20220306A0264S00.
[5] 蒋林涛.数据网的现状及发展方向[J].电信科学,2019,35(8):1-15.
[6] IETF.IPv6 segment routing header (SRH):RFC 8754[S/OL].[2022-01-04].https://www.ietf.org/.
[7] IETF.Segment routing over IPv6(SRv6)network programming;RFC 8986[S/OL].[2022-01-04].https://www.ietf.org/.
[8] IETF.In-situ flow information telemetry framework;draft-song-opsawg-iftf-fr-amework-00[EB/OL].[2022-01-04].https://tools.ietf.org/id/draft-song-opsawg-iftf-framework-09.html.
[9] IETF.Alternate-marking method for passive and hybrid performance monitoring;RFC 8321[S/OL].[2022-01-04].https://www.ietf.org/.
[10] IETF.BGP SR Policy for IFIT;draft-ietf-idr-sr-policy-iftf[EB/OL].[2022-01-04].https://www.ietf.org/.
[11] IETF.Usecase/design consideration draft-cheng-spring-ipv6-msr-design-consideration[EB/OL].[2022-01-04].https://www.ietf.org/.
[12] IETF.BIER IPv6 Requirements;draft-ietf-bier-ipv6-requirements-01[EB/OL].[2022-01-04].https://datatracker.ietf.org/doc/draft-ietf-bier-ipv6-requirements/01/.
[13] IETF.Source segment;draft-xl-msr6-source-segment[EB/OL].[2022-01-04].https://www.ietf.org/.
[14] IETF.MSR6 Traffic Engineering;draft-geng-msr6-traffic-engineering-00[EB/OL].[2022-01-04].https://www.ietf.org/.
[15] 王晨曦,古锐,肖亚群,等.基于“IPv6+”的智能IP网络方案[J].电信科学,2020,36(8):15.
[16] 我们离“算力网络”的星辰大海还有多远?[EB/OL].[2021-11-12].https://baijiahao.baidu.com/s?id=1716148246895146673&wfr=spider&for=pc.
[17] 解读下一代网络:算力网络正从理想照进现实[EB/OL].[2020-11-20].https://bbs.huaweicloud.com/blogs/211409.
[18] 曹畅,佟恬,屠礼彪.引领“IPv6+”创新之路 中国联通在多地开展实践[J].通信世界,2022(4):3.
[19] 张惠卿,沈金龙,钱辉兵,等.下一代IP网络技术——IPv6的研究及其演进方案[J].中国数据通信,2002,4(11):5.
[20] 魏旻,胡港慧,王平,等.基于IPv6的工业互联网关键技术及其标准化研究[J].信息技术与标准化,2022(3):6.
[21] 李建昊,张宏科.基于双栈与隧道结合技术建构IPv6网络[J].北京交通大学学报,2004,28(5):58-61.

作者简介:

秦壮壮,高级工程师,主要研究方向为IPv6、IP城域网和IP承载网的网络架构与技术演进、SDN和项目管理等;赵金水,高级工程师,主要研究方向为IPv6、IP城域网及接入、IPTV、网络安全的技术演进与系统管理;屠礼彪,高级工程师,主要研究方向为IPv6、IP城域网、智能城域网、IPRAN的网络架构与技术演进、规划及建设方案管理等;王笙,高级工程师,主要研究方向为IPv6、智能城域网和光传送网络的网络架构与技术演进、项目管理等;刘敏,高级工程师,主要研究方向为IP城域网、IP承载网及骨干承载网的建设与项目管理,IPv6、云汇接等技术发展和演进等。