

算力网络关键技术及发展挑战分析

曹畅 唐雄燕

(中国联合网络通信有限公司研究院,北京 100048)

摘要: 分析了融合计算、存储、传送资源的新一代网络基础设施算力网络的提出背景,结合网络云化、协议创新和网络能力开放论述了算力网络的关键技术和演进趋势,提出了算力网络未来发展中面临的挑战,给出了若干促进算力网络发展的建议。

关键词: 算力网络; 云网融合; IPv6; 可编程芯片

中图分类号: TN929.5 文献标识码: A

引用格式: 曹畅,唐雄燕. 算力网络关键技术及发展挑战分析[J]. 信息通信技术与政策,2021,47(3):6-11.

doi: 10.12267/j.issn.2096-5931.2021.03.002

0 引言

新一代信息网络正在从以信息传递为核心的网络基础设施,向融合计算、存储、传送资源的智能化云网基础设施发生转变,算力网络是为应对这种转变而提出的新型网络架构^[1-2]。算力网络基于无处不在的网络连接,将动态分布的计算与存储资源互联,通过网络、存储、算力等多维度资源的统一协同调度,使海量的应用能够按需、实时调用泛在分布的计算资源,实现连接和算力在网络的全局优化,提供一致的用户体验。作为一项新兴技术,算力网络的发展趋势、技术特征以及所面临的发展挑战正在逐步显现,算力网络的健康、可持续发展对我国建设网络强国,构建数字基础设施建设体系意义重大。

1 计算与网络融合发展趋势分析

数字化、智能化正在加快推动全球 ICT 产业的创新。世界各主要经济体国家纷纷通过国家战略来抢占数字经济产业链的制高点,如美国的《美国国家网络战略》、德国的《高技术战略 2025》、日本的《日本制造业白皮书》等。此外,数字化、智能化还催生了海量的场景和应用,5G 特性将带动 VR、AR、云游戏等大带宽、低时延业务走向商业化,同时奠定 Cloud+X 发展基

础。从流量看,未来移动数据流量会继续剧增。据爱立信预测,2025 年全球移动数据流量将较 2019 年增加 5 倍,其中 45% 的流量将由 5G 网络承载^[3]。网络带宽的扩大与业务类型的不断拓展将增加网络中的视频流量,以国内电信运营商中国联通为例,2020 年 5 月其视频流量占比已达 72%,预计到 2025 年占比将增至 76%。

新应用带来的爆炸式数据增长将重塑 ICT 产业格局,ICT 技术发展将向着泛在联接与泛在计算紧密结合的方向演进,促使满足多样化计算需求的融合计算架构出现,以及云一边一端结合的泛在计算模式兴起。伴随着我国数字经济的蓬勃发展,新一代信息技术间的融合效应渐显,“5G+云+AI”将成为推动我国数字经济持续发展的重要引擎,基于上述背景,我国率先提出了“算力网络”的概念。算力网络需要网络和计算高度协同,将计算单元和计算能力嵌入网络,实现云、网、边、端、业的高效协同,提高计算资源利用率。在算力网络中,用户无需关心网络中计算资源的位置和部署状态,而只需关注自身获得的服务,并通过网络和计算协同调度保证体验一致^[4-5]。

2 算力网络发展情况

目前来看,我国信息通信行业对“5G+云+AI”的探

索处于世界领先地位。据 IDC 预测,到 2023 年,数字经济产值将占到中国 GDP 的 67%,超过全球平均水平。这些都带动了全网的算力密集分布、快速下沉并逐步实现联网服务。截至目前,算力网络的愿景已在业界得到广泛的认可;算力网络在标准制定、生态建设、试验验证等领域均取得了一定进展;算力网络作为我国的一项原创成果,开始走向国际舞台。在标准制定方面,中国移动、中国电信与中国联通分别在国际电信联盟 (ITU-T) SG11 与 SG13 工作组立项了 Y. CPN^[6]、Y. CAN 和 Q. CPN^[7] 等系列标准,并在互联网工程任务组 (IETF) 开展了 Computing First Network Framework^[8] 等系列标准的研究;华为联合国内运营商在欧洲电信标准化协会 (ETSI) 和世界宽带论坛 (BBF) 启动了包括 NW1、城域算网在内的多个项目;中国通信标准化协会 (CCSA) 正有序开展“算力网络需求与架构”以及“算力感知网络关键技术研究”两项工作。面向未来 6G 时代,国内 IMT-2030 6G 网络工作组已将算力网络作为研究课题之一,并开展算力网络与 6G 通信技术的融合研究。在生态建设方面,国内未来数据通信研究的主要组织——网络 5.0 产业联盟专门成立了“算力网络特设工作组”。2019 年年底,中国联通、中国移动和边缘计算网络产业联盟 (ECNI) 均发布了算力网络领域相关白皮书,进一步阐述了算网融合的重要观点;2020 年年底,中国联通率先成立“中国联通算力网络产业技术联盟”,作为首个运营商牵头的算力网络研究组织,结合自身业务发展,对相关先进网络协议的制定提出了明确需求。在试验验证方面,2019 年中国电信与中国移动均已完成算力网络领域的实验室原型验证,并在 GSMA 巴塞罗那展、ITU-T 和全球网络技术大会 (GNTC) 相关展会上发布成果。2020 年年底,中国联通在江苏南京开通了国内首个集成开放网络设备、算力服务平台和 AI 应用的一体化试验局。

2020 年 4 月,国家发展和改革委员会首次对新基建的具体含义进行了阐述,在信息基础设施部分,提出构建以数据中心、智能计算中心为代表的算力基础设施。2020 年 12 月,国家发展和改革委员会在《关于加快构建全国一体化大数据中心协同创新体系的指导意见》中提出“布局大数据中心国家枢纽节点,形成全国算力枢纽体系”的具体要求。从政策的导向来看,国家

给予了算力提供者、网络运营者、服务提供者和使用者等不同角色引入多方参与的空间,同时也给以算力网络技术为基础的转—算—存主体分离、联合服务的新商业模式提供了宝贵的尝试空间^[9]。

3 算力网络关键技术分析

从算力网络所倡导的技术理念中可以看出,算网一体是结合 5G、泛在计算与 AI 的发展,在云网拉通和协同基础上的下一个网络发展阶段,即云网融合 2.0 阶段。云网融合 2.0 阶段在继承云网融合 1.0 阶段工作的基础上,强调结合未来业务形态的变化,在云、网、芯 3 个层面持续推进研发,实现应用部署匹配计算、网络转发感知计算、芯片能力增强计算,服务算力网络时代云游戏、千人千面直播、自动驾驶、智能安防与工业机器视觉等新业态,其关键技术可归类为网络编排、网络承载和网络转发 3 个方面^[10]。

3.1 网络编排技术

算力网络是融合计算、存储、传送资源的智能化新型网络,通过全面引入云原生技术,实现业务逻辑和底层资源的完全解耦。需通过打造面向服务的容器编排调度能力,实现服务编排向算网资源的能力开放。同时,可结合底层基础设施的资源调度管理能力,对于数据中心内的异构计算资源、存储资源和网络资源进行有效管理,实现对泛在计算能力的统一纳管和去中心化的算力交易,构建一个统一的服务平台。

基于云原生的算力建模与服务编排如图 1 所示,可见,算力是设备/平台处理、运行业务的关键核心能力。在算力网络中,算力的提供方不再是专有的某个数据中心或集群,而是将云边端这种泛在的算力通过网络化的方式连接在一起,实现算力的高效共享^[11]。因此,算力网络中的算力资源将是泛在化的、异构化的。对于异构算力资源,算力网络架构采用基于“K8S+轻量化 K8S”的两级联动的架构来实现统一的算力资源调度纳管。K8S 采用中心的资源调度统一平台对于整体的基础资源进行统一管理和集群管理,而轻量化 K8S 集群主要是作为边缘侧的资源调度平台对于边缘计算集群进行调度和管理。

泛在算力资源的统一建模度量是算力调度的基础。针对泛在的算力资源,通过模型函数将不同类型的算力资源映射到统一的量纲维度,形成业务层可理

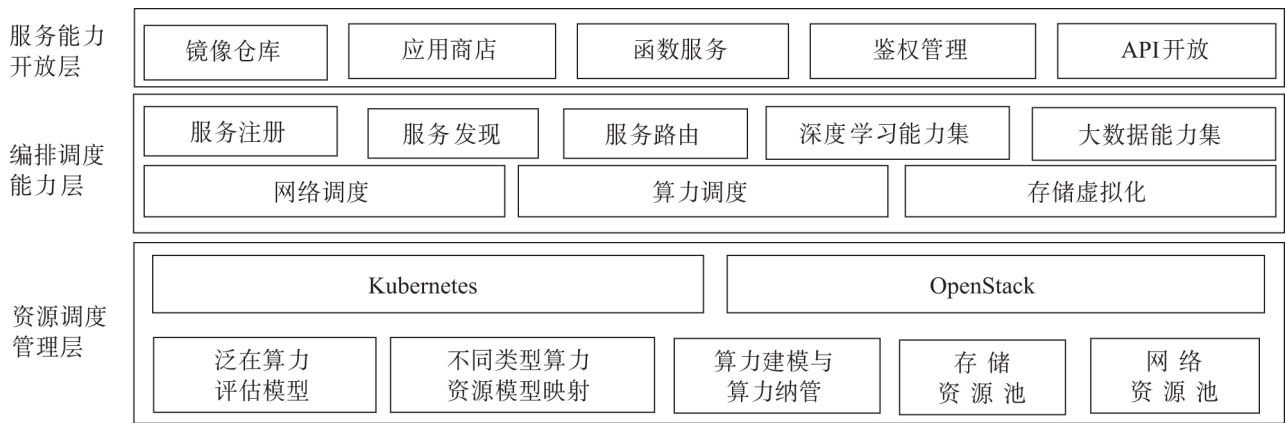


图 1 基于云原生的算力建模与服务编排示例

解、可阅读的零散算力资源池,为算力网络的资源匹配调度提供基础保障,将业务运行所需的算力需求按照一定分级标准划分为多个等级。这样可为算力提供者设计业务套餐时提供参考,也可为算力平台设计者根据所需运行的业务做平台算力的选型设计提供依据。

3.2 网络承载技术

“IPv6+”是面向 5G 和云时代的智能 IP 网络,可以满足算力网络灵活组网、优化用户体验按需服务等需求。算力承载网以 IPv6+和 SRv6 技术为底座,在网络切片能力的基础上,引入网络感知技术,解决当前网络难以感知业务需求、算力和服务难以良好匹配的问题。在算力网络中,业务网关进一步下沉,并通过算力网关将南北向流量提前转化为东西向流量,同时利用 IPv6 可扩展头丰富的可编程空间,开展 IPv6+ 网络新技术

包括但不限于网络切片、随路网络监测(in-situ Flow Information Telemetry , iFIT)、业务链(Service Function Chaining , SFC) 等和新应用开发,实现城域算力基础设施互联,通过业务的部署和资源调整来保证应用的服务级别协议(Service Level Agreement , SLA) 要求,以此提供业务链服务,具体如图 2 所示。当前,CCSA TC3 正在开展算力网络需求与架构的研究工作,思路是通过对网络、存储、算力等多维度资源的统一管理和协同调度,实现连接和算力在网络的全局优化,技术实现上可分为集中式方案、分布式方案、混合式方案等多种技术路线。除基于数据中心软件定义网络(Software Defined Network , SDN) 集中调度的方案,还可基于电信运营商承载网分布式控制能力,结合承载网网元自身控制协议扩展,复用现有 IP 网络控制平面分布式协

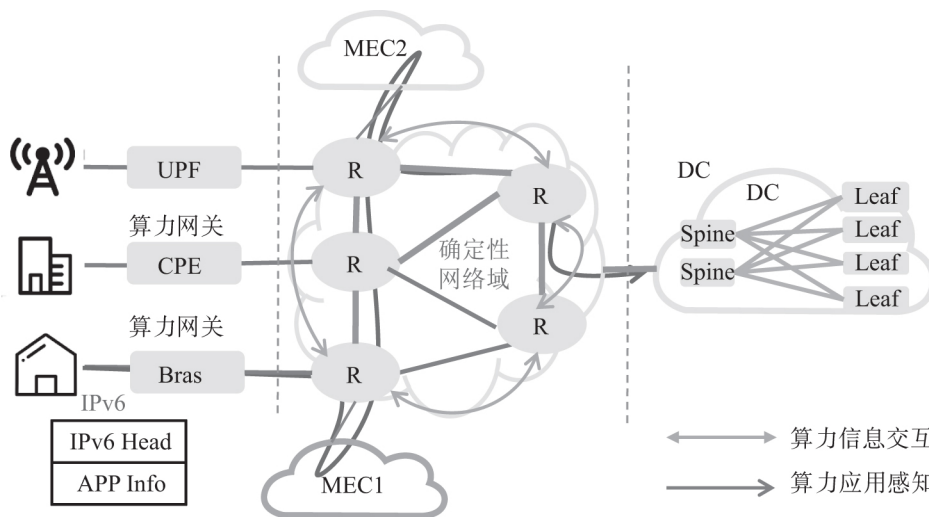


图 2 算力网络资源感知和信息交互示例

议的方式实现算力信息的分发与基于算力寻址的路由,同时综合考虑实时的网络和计算资源状况,将不同的应用调度到合适的计算节点进行处理。

3.3 网络转发技术

传统的网络设备采用转控一体的工作模式,其转发芯片的功能相对固化,“紧耦合”的网络设备难以支撑算力网络对设备灵活性及可编程性的需求。新一代高性能可编程数据包处理芯片加上 P4/NPL 等编程语言的出现,让网络所有者、工程师、架构师及管理员可以自上而下地定义数据包的完整处理流程,除了可帮助算力网络实现最适合其自身需求的具体网络行为外,其中可编程芯片还能使芯片供应商专注于设计并改进那些可重用的数据包处理架构和基本模块,而不必纠结于特定协议里错综复杂的细节和异常行为。因此,可编程芯片技术的产生,为算力网络转发面提供了相应的技术支撑。

4 算力网络发展挑战与建议

4.1 算力资源的感知与度量

算力网络在工程实际应用中首先面临的是算力的感知与度量,进而才能实现对算力的编排并合理快速匹配业务需求。目前,如何感知算力、通过有效建模形成统一度量的算力资源,并能够合理编排以满足业务需求,是算力网络研究的重点和难点。

随着 5G、人工智能等技术的发展,算力网络中的算力提供方不再是专有的某个数据中心或计算集群,而是云边端这种泛在化的算力,并通过网络连接在一起,实现算力资源的高效共享。因此,算力网络中的算力资源将是泛在化的、异构化的。目前,市面上不同厂家的计算芯片类型形式各异,如 GPU、ASIC,以及近年出现的 NPU、TPU 等,这些芯片的功能和适用场景各有侧重,如何准确感知这些异构的泛在芯片的算力大小、不同芯片所适合的业务类型及其在网络中的位置,并且对其进行有效纳管、监督是目前的主要挑战。

针对上述挑战,建议结合算力需求量化与建模研究,积极推动相关国际国内标准化工作,通过标准化的模型函数将不同类型的算力资源映射到统一的量纲维度,形成业务层可理解、可阅读的零散算力资源池。另外,对于业务运行,不光要有足够的算力,也需要配套的存储能力、网络能力,甚至还可能需要编解码能力、

吞吐能力等来联合保障用户的业务体验。建议从微服务的角度来衡量算力,对相应的资源调度分配原则进行标准化,降低算力网络中业务和应用部署的复杂度,简化业务管理流程和机制^[12]。

4.2 集中与分布的协同控制

算力网络控制方案的实现有集中式和分布式两种。集中式方案是在基于数据中心 SDN 集中调度方案的基础上,由云数据中心向城域网扩展,与边缘云相连接,通过集中式的 SDN 控制器和网络功能虚拟化编排器管理和协调功能(NFVO MANO)实现中心云及边缘云间的算力网络的统一管理和协同调度。分布式控制方案即基于电信运营商承载网分布式控制能力,结合承载网网元自身控制协议扩展,复用现有 IP 网络控制平面的方式实现算力信息的分发与基于算力寻址的路由,同时综合考虑实时的网络和计算资源状况,将不同的应用调度到合适的计算节点进行处理,实现连接和算力在网络的全局优化^[13]。

对比集中式控制与分布式控制两种方案,前者能够做到算力节点的路由可达,配置通过集中式的 SDN 控制器可快速实现,但该方案的问题是计算节点无法快速与网络属性联动,也较难与运营商基础网络联动;后者能够充分调动承载网中 IP 路由器节点的控制能力,应用可以感知路径中沿途的所有节点的服务质量,但需要网络根据具体的业务需求选择边界网关协议(Border Gateway Protocol, BGP) 扩展的种类和形式,实现比较复杂,也尚未标准化。

针对上述挑战,建议大力推进算力感知与端管云协同的技术研发,例如“IPv6+”系列技术研发。“IPv6+”拉通端管云以实现统一的网络配置,可以满足云网融合的灵活组网、业务快速开通、确定性传输、优化用户体验按需服务等需求,通过 IPv6 协议与扩展,可以使多方、异构的资源整合在一起,解决云和网的灵活对接、云网资源的统一管控和资源利用的整体最优化。同时,建议在“IPv6+”协议基本功能具备之后,研究云服务应用感知、算力资源及时调用与网络能力开放之间的协调机制,以便更好地推进云网融合,促进算力的端管云拉通^[14]。

4.3 计算和网络的联合布局优化

从过去来看,计算和网络两大产业虽互有融合和促进,但总体上还处于分别发展、独自规划的阶段。

5G 时代的到来,对计算和网络的联合布局优化提出了必然的要求。其一,单芯片、单设备的计算能力遇到了制造工艺、多核集成数量等方面的瓶颈,这就要求多芯片、多算力设施的联合服务。其二,5G 核心网的云化部署使得边缘计算成为了可能,边缘计算要求计算的单元贴近用户,网络的服务质量成为评价边缘计算基础能力的重要标准。其三,随着 AI 识别、大视频、科学计算等新业务的发展,算力类型在 CPU 通用计算的基础上,不断向 GPU、ASIC 等专用类型扩展,需结合用户快速接入计算服务的要求,计算节点在网络中的布局也需要结合网络情况和业务需求综合考虑^[15-16]。

针对上述挑战,建议加强顶层设计,通过“以算联网,以网促算”的方式进行计算和网络的联合布局优化,并通过计算成网弥补我国计算芯片单体的自主可控短板。具体来说,需要加强计算处理单元和网络控制系统双方的开放性,以便更加快速、便捷地响应对方的需求。同时,建议运营商在 5G 核心网、承载网部署过程中,一方面做强边缘计算,形成云一边一端多级算力有效协同和分担的局面;另一方面提升网络承载能力,综合建网成本和算力传递质量需求进行骨干和城域网承载网的设计和建设。

5 结束语

数字经济成为“十四五”的重要创新增长引擎,国家把“网络强国、数字中国”作为“十四五”新发展阶段的重要战略进行了系列部署。信息通信行业深处“网络强国、数字中国”建设的宏观战略基点,立足数字产业化、产业数字化的时代风口,应全力围绕数字经济“新需求”创造“新供给”,积极探索构建算力网络、提供算力服务的方式与途径,打造高品质网络优势,携手产业伙伴与广大用户,共创数字经济的美好未来。

参考文献

- [1] 何涛,曹畅,唐雄燕,等. 面向 6G 需求的算力网络技术[J]. 移动通信,2020,44(6):131-135.
- [2] 林德平,彭涛,刘春平. 6G 愿景需求、网络架构和关键技术展望[J]. 信息通信技术与政策,2021,47(1):82-89.
- [3] 爱立信. 爱立信移动市场报告[R],2020.
- [4] 中国联合网络通信集团有限公司. 中国联通算力网络

白皮书[R],2019.

- [5] 程强,刘姿杉. 电信网络智能化发展现状与未来展望[J]. 信息通信技术与政策,2020,46(9):16-22.
- [6] ITU-T. Draft recommendation ITU-T Y. CPN-arch: framework and architecture of computing power network[S],2020.
- [7] ITU-T. Draft recommendation ITU-T Q. CPN: signaling requirement of computing power network[S],2019.
- [8] IETF. Framework of Compute First Networking (CFN) draft-li-rtgwg-cfn-framework-00[S],2019.
- [9] 李勇坚. 全球“新基建”热潮的理论解析[J]. 信息通信技术与政策,2020,46(9):8-15.
- [10] 中国联合网络通信集团有限公司. 算力网络架构与技术体系白皮书[R],2020.
- [11] 中国移动研究院. 算力感知网络技术白皮书[R],2019.
- [12] 中国通信学会. 算力网络前沿报告[R],2020.
- [13] 曹畅,张帅,刘莹,等. 基于通信云和承载网协同的算力网络编排技术[J]. 电信科学,2020,36(7):55-62.
- [14] 网络 5.0 产业和技术创新联盟. 网络 5.0 技术白皮书[R],2019.
- [15] 全球移动通信系统协会(GSMA). 5G 时代的边缘计算:中国的技术和市场发展[R],2020.
- [16] 边缘计算网络产业联盟(ECNI). 运营商边缘计算网络技术白皮书[R],2019.

作者简介:

- 曹畅** 中国联合网络通信有限公司研究院未来网络研究部高级专家,智能云网技术研究室主任,第七届中国通信学会信息通信网络技术委员会委员,中国通信标准化协会 CCSA “网络 5.0 技术标准推进委员会”架构组副组长,边缘计算网络基础设施联合工作组(ECNI)技术规范组组长,主要专业领域为 IP 网络宽带通信、SDN/NFV、新一代网络编排技术等,获中国联通科技进步奖 2 项,已发表论文 30 余篇,专利授权 10 余项
- 唐雄燕** 中国联合网络通信有限公司研究院首席科学家,北京邮电大学兼职教授、博士生导师,工

业和信息化部通信科技委委员兼传送与接入
专家咨询组副组长,北京通信学会副理事长,
中国通信学会理事兼信息通信网络技术委员
会副主任,中国光学工程学会常务理事兼光
通信与信息网络专家委员会主任,国际开放

网络基金会(ONF)董事,20 余年电信新技术
新业务研发与技术管理经验,主要研究领域
为宽带通信、光纤传输、互联网/物联网、
SDN/NFV 与新一代网络等

Analysis of key technologies and development challenges of computing power network

CAO Chang ,TANG Xiongyan

(China Unicom Research Institute ,Beijing 100048 , China)

Abstract: This paper analyzes the background of the new generation of network infrastructure computing power network which integrates computing , storage and transmission resources , then discusses the key technologies and evolution trend of computing power network in combination with network cloudification , protocol innovation and network capacity opening , and points out the challenges in the future development of computer power network , finally puts forward some suggestions to promote the development of computer power network.

Keywords: computing power network; cloud network integration; IPv6; programmable chip

(收稿日期: 2021-02-18)