

大数据与决策研究

(政策与技术跟踪专题)

2022年第23期(总第133期)

广西壮族自治区信息中心
广西壮族自治区大数据研究院

2022年7月26日

编者按：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。算力网络正是为应对这种转变而提出的新型网络架构。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。本期主要介绍算力网络相关技术情况。

本期要目

算力网络的概念与组网架构

算力网络核心技术体系

算力网络的典型应用场景

算力网络的概念与组网架构

算力网络是电信运营商为应对云网融合向算网一体转变而提出的新型网络架构，是实现算网一体的重要技术抓手。图 1 所示算力网络组网架构图，将算力网络按照功能区分为四个域，分别为接入网络域、算网网关域、算网承载域和数据中心域：

接入网络域：以“极”为核心特征，针对各种用户接入网络的南北向流量，实现极致化的大带宽、低延时、广连接等通信指标；

算网网关域：以“柔”为核心特征，结合各种接入业务的具体特征，面向固移网络融合承载、控制转发面分离、转发面下沉的演进需求，实现算网网关柔性、弹性、低成本部署；

算网承载域：以“智”为核心特征，算网承载域是算力网络的核心，需要承载网结合 SRv6¹、切片²、APN6³、ROADM⁴等新技术满足东西向流量的承载需求，实现业务在多云之间的智能调度，通过引入算力感知、业务感知、确定性服务等能力，结合运营商城域、骨干等多级架构实现；

数据中心域：以“简”为核心特征，面向数据中心内云

¹ 即基于 IPv6 的段路由，其采用现有的 IPv6 转发技术，通过灵活的 IPv6 扩展头，实现网络可编程。

² 网络切片是一种按需组网的方式，可以让运营商在统一的基础设施上分离出多个虚拟的端到端网络，每个网络切片从无线接入网到承载网再到核心网上进行逻辑隔离，以适配各种各样类型的应用。

³ 即基于“IPv6+”的应用感知网络，通过将应用的需求信息封装在数据分组中，使网络能感知应用及其需求，便于网络进行流量调度和资源调整。

⁴ 即可重构光分插复用器，是一种使用在密集波分复用（DWDM）系统中的器件或设备，其作用是通过远程的重新配置，可以动态上路或下路业务波长。

服务的承载需求，实现数据中心内部网络架构的简化和高效、无损传输。

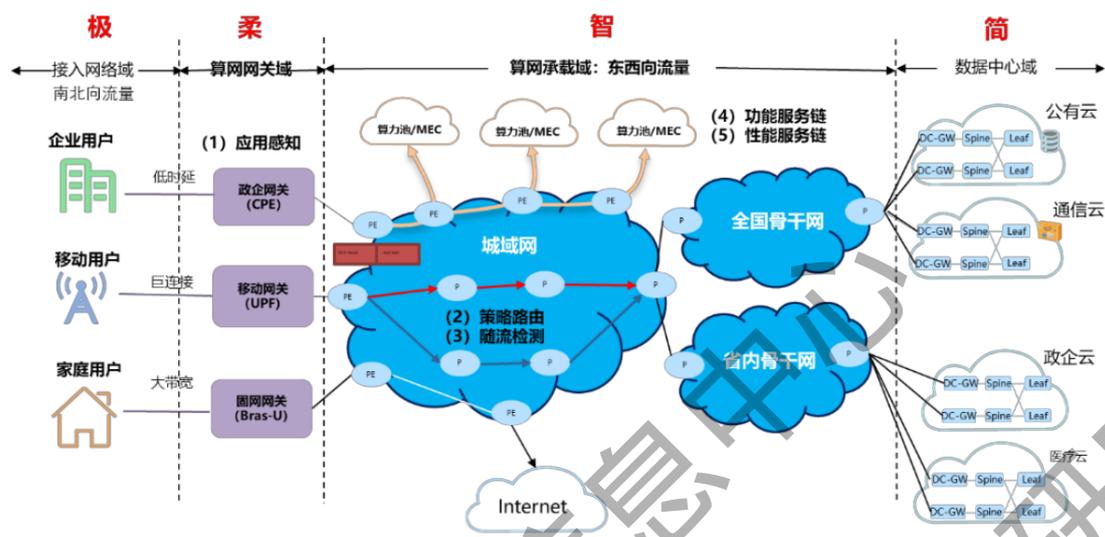


图 1 算力网络组网架构

(来源:《2021 中国联通算力网络实践案例》)

算力网络核心技术体系

研究者基于算力网络三层架构和技术图谱，在提出算力原生、算力度量、算力路由等新技术理念的基础之上，进一步对算力网络的关键技术进行了体系化的梳理和深度挖掘，聚焦三十二大核心技术，与十大发展方向，三层架构进行了关联映射，如图 2 所示。

算力网络的发展需要在推动单点创新技术成熟的基础上，穿通多种关联技术，形成针对纵向技术栈的端到端解决方案，实现算力网络技术簇的有效连接和共同成熟，避免产生“技术孤岛”。为此，研究者进一步梳理各核心技术之间的关联联系，具体如下：

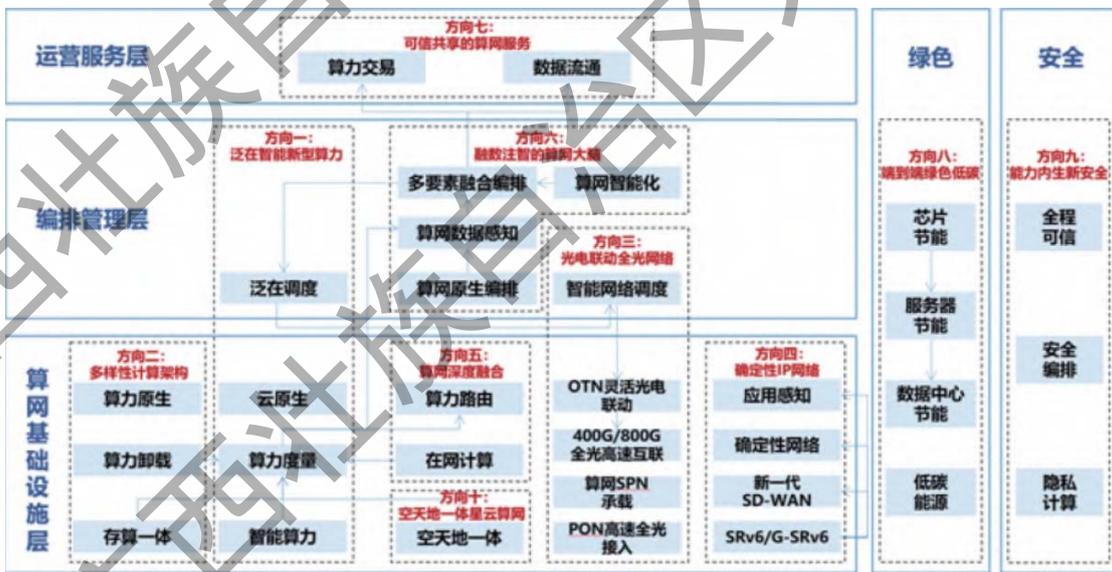


图 2 算力网络核心技术体系

算力方面：存算一体、智能算力、在网计算、算力卸载、空天地一体等算力技术通过层次化的算力度量评价评测体

系，形成标准可量化的能力度量模型，一方面使能算力路由实现分布式算力调度，一方面赋能算网数据感知，实现算力节点的有效算力事实上报。

网络方面：传送网高速全光接入⁵和算网 SPN⁶提供灵活敏捷的接入方式，全光高速互相结合光电联动技术实现传输资源的动态调度和实时响应。同时，智能网络调度向上对接泛在调度，实现传送网基于业务需求的灵活全光调度，向下实现传送网各专业拉通。承载网以 SRv6/G-SRv6 产品和切片功能，并基于 SRv6/G-SRv6 的可编程特性构建应用感知、确定性保障的差异化网络服务。

编排管理和运营服务方面：算网数据感知实施捕捉算网业务和资源的动态特性，算网原生编排提供算网服务化、意图化接口，算网智能化构建算网自智能能力，三者共同支撑多要素融合编排。多要素融合编排向下对接泛在调度，实现算网资源和应用的跨域拉通和部署，向上为算力交易和数据流通新型算网服务的构建提供能力支撑，实现多技术要素融合能力供给。

绿色安全方面：结合低碳能源技术，从芯片节能、服务器节能、数据中心节能，由点及面，端到端的节能技术逐层递进；隐私计算、安全编排、全程可信技术共同打造一体化安全内生防护机制。

（来源：《2022 年算力网络技术白皮书》）

⁵ PON 即光接入网技术，是一种典型的无源光纤网络，是指（光配线网中）不含有任何电子器件及电子电源，ODN 全部由光分路器（Splitter）等无源器件组成，不需要贵重的有源电子设备。

⁶ 即切片分组网，是 5G 网络切片中的关键技术，用其来支撑下一代网络架构，带宽，流量模式，切片，延迟及时间同步。

算力网络的典型应用场景

算力网络与 5G 和 AI 的技术成熟度存在密切关系，最先成熟的应用场景将主要集中在借助视频和图像 AI 技术实现自动识别和辅助决策系统。这些领域共同的需求体现在：

- 无线化：包括 MR/AR/摄像头/机器人等多种灵活的无线接入方式；

- 大带宽：高清视频传输需要的网络；

- 确定时延：工业现场毫秒级的确定时延；

- 高算力：基于图像分析的 AI 推理能力、图像渲染需要的计算能力；

- 碎片化：对算力的需求来自各行各业零散需求，存在突发和碎片化的特点；

- 可信：由于要涉及数据的处理，企业对算力处理有可信性，价值可量化、可交易的需求。

目前已经出现的场景，覆盖了运营商 ToB 和 ToC 领域。

一、运营商 ToB 的“5G 园区+AI”场景

企业建设 5G 园区的主要驱动力来自视频类业务的驱动力，包含监控、智能生产、园区自动驾驶等需要更大带宽和更高实时处理的业务需求。这类业务对边缘智能方案有几个约束：

- 数据隐私要求：原始数据不上公有云；

●节省带宽成本：不显著增加专线和数据中心的成本，部分原始数据本地分析处理，按需存储在数据中心；

●极限边缘部署：室外或 IT 机房场景并存，对噪音、功耗有要求，需要支持无风扇自然散热、更宽的环境温度和低功耗；

●低时延要求：企业内一些与安全、生产相关的 AI 应用，需要 ms 级的低时延。

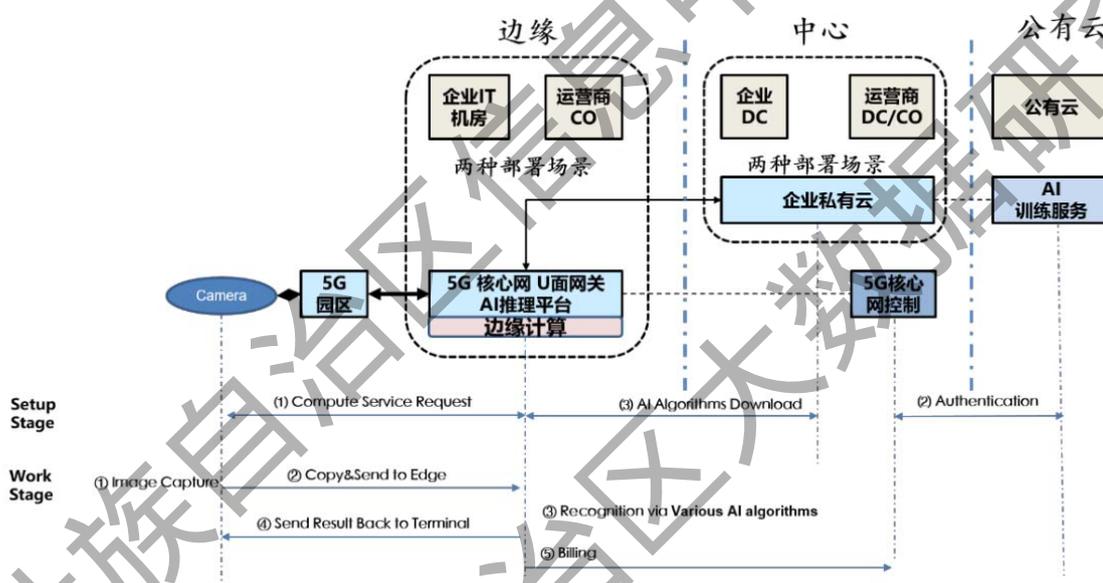


图3 5G 园区+AI 应用场景

在企业边缘，需要在有限的功耗和空间约束下，支持核心网 U 面功能、支持视频 AI 处理，并要考虑将多个边缘的算力共享，组成资源池，最大限度地提升算力的效率。公有云上可以提供一些企业 AI 训练的赋能服务。

二、运营商 ToC 的“5G+Cloud X”场景

带宽和时延是 Cloud X 业务体验的最大挑战，其中 5G 空口和核心网处理时延相对稳定，传输时延受网络环境（如

拥塞)，传输距离的影响较大。以 Cloud VR 为例，为保障良好体验，需要端到端时延小于 40ms，要保障极致体验，需要端到端时延小于 20ms，而服务器和终端处理时延已经超过 15ms，要保障极致体验，留给网络“5G+传输+MEC”的总时延只有不到 5ms。

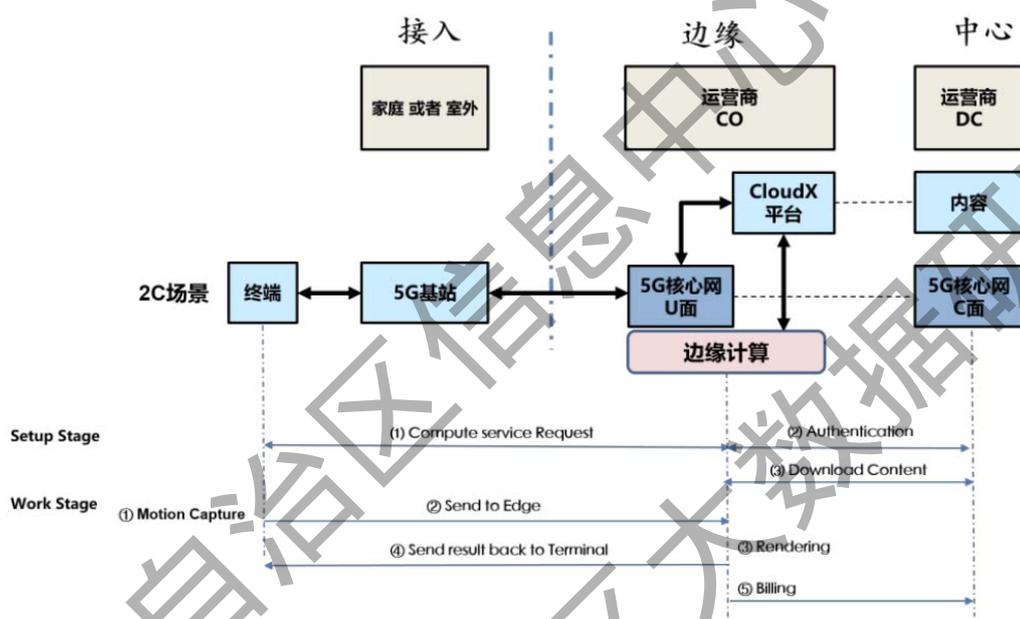


图 4 5G+Cloud X 应用场景

为保障业务体验，满足苛刻的时延要求，需要考虑 C/U 分离的核心网架构，U 面下移到运营商边缘，靠近用户提供服务。Cloud X 业务主要需要借助 GPU 的算力，为降低每用户服务成本，需要提高算力的使用效率。为保障业务体验，需要减少网络时延抖动的发生，需要业务和网络之间增加感知和预测能力，在业务发放时，根据网络状态主动做出均衡。

三、算力开放, 运营商提供可交易的算力通证

智能社会里，数以万计的互联网公司流量的获取者、算力的需求者，而运营商拥有大量的边缘机房和网络云化后

边缘算力的优势,是算力的拥有者和流量的引导者。为了适应智能社会的发展,一种动态、快速匹配上述供需关系的新型商业模式是:互联网公司/第三方从运营商按需调用边缘算力,流量获取来自运营商的算力网络,从而节省大量边缘计算重资产投资和流量获取成本;运营商释放边缘机房、网络云化优势,变现“网络连接及其相关的数字资产(计算和存储)”。智能社会带来的这种基于共享经济的新商业模式,可以结合图5所示的技术方案加以描述:

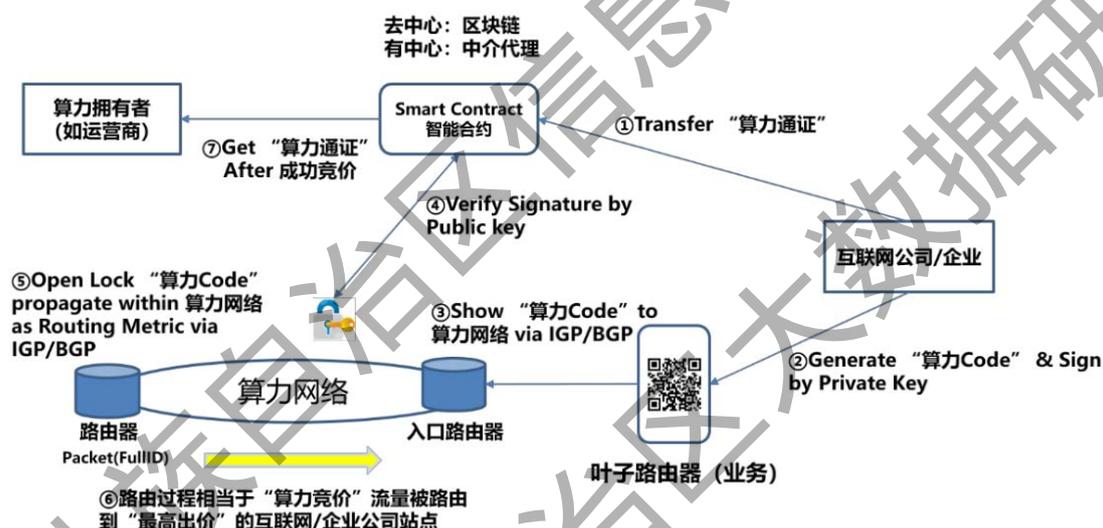


图5 通过算力通证进行算力交易应用场景

需要说明的是,“算力通证”与运营商计算站点可提供的计算能力是等价,即“算力通证”是计算站点可提供的计算能力的一种商业抽象,而“算力 Code”是计算站点可提供的计算能力的一种技术抽象,“算力通证”以“算力 Code”的技术形式在算力网络中传递。

(来源:《中国联通算力网络白皮书》)

广西壮族自治区信息中心
广西壮族自治区大数据研究院

编辑部地址：南宁市体强路 18 号广西信息中心 1412 号房

联系电话：0771-6113592

电子邮箱：dsjyjs@gxi.gov.cn

网 址：<http://gxxxxx.gxzf.gov.cn/>



扫描二维码获取
更多决策参考信息