

# 适应算力需求的承载网络架构研究

## Research on Bearer Network Architecture Adapting to Computing Power Demands

张纬卿(上海邮电设计咨询研究院有限公司,上海 200092)

Zhang Weiqing(Shanghai Posts&Telecommunications Designing Consulting Institute Co.,Ltd.,Shanghai 200092,China)

### 摘要:

简要归纳了算力业务的典型场景,通过分析算力驱动的业务核心要素变化趋势,从网络拓扑、组网灵活性、网络容量、网络无损性能、智能调度能力等方面论述了面向算力的承载网络演进策略。在此基础上,提出了算力网络整体架构模型、算力承载网络的架构组成及关键技术。

### 关键词:

算力承载网络;算力路由;确定性网络;400G;细粒度 OTN;SRv6

doi:10.12045/j.issn.1007-3043.2026.02.010

文章编号:1007-3043(2026)02-0050-06

中图分类号:TN913

文献标识码:A

开放科学(资源服务)标识码(OSID):



### Abstract:

It briefly summarizes typical scenarios of Computability Business. By analyzing the changing trend of business core elements driven by computing power, it discusses the evolution strategy for bearer networks oriented towards computing power in terms of network topology, flexibility, capacity, lossless performance, and intelligent scheduling capabilities, etc. On this basis, the overall architecture model of the computing power network, the architecture composition and key technologies of the computing power bearing network are proposed.

### Keywords:

Computing power bearer network; Computing power routing; Deterministic network; 400G; Fine-grained OTN; SRv6

引用格式:张纬卿. 适应算力需求的承载网络架构研究[J]. 邮电设计技术,2026(2):50-55.

## 0 引言

随着多模态大型 AI 模型的技术突破,大规模智算正快速迈向“超万卡、超万亿参数”时代,各国政府和企业都在加速推动智算基础设施的布局和升级。2024 年 1 月国家发展改革委等部门下发了《关于深入实施“东数西算”工程加快构建全国一体化算力网的实施意见》,国家正致力于推动算力成为数字经济时代的新型生产力,建设支撑数字经济高质量发展的关键基础设施——算力承载网络。

在此背景下,现有的承载网络面临诸多能力与业务趋势不相适配的挑战,如拓扑弹性和可扩展性不足,系统容量和接口与算力业务发展需求不相适应、业务感知能力缺乏、智能动态路由能力偏弱等。如何高效稳固地构建新一代算力承载网络,以网强算,夯实数字基础设施的“底座”,成为承载网络亟待解决的新课题和新挑战。

## 1 算力业务形态及演进趋势分析

### 1.1 算力业务的典型场景

根据端到端流向、流量调度的距离、复杂度等情况的不同,节点间算力业务的主要场景可归纳如下。

收稿日期:2025-12-30

### 1.1.1 智算DCN(分布式)连接场景

智算DCN(分布式)连接场景主要包括智算类型的跨域连训(参数面拉远)等。其典型的业务特征为10T~100T级带宽、极低丢包和延时等。目前的研究热点主要集中在城域、一体化片区等面向万卡集群、万亿级大模型的跨域连训短距离场景、基于全国一体化算力网的“东数西算”场景以及将大量中小算力节点整合为一体化算力池的场景等。

### 1.1.2 DCI高可靠互联场景

DCI高可靠互联场景主要包括通算类型的多AZ集群节点间的DCI大带宽、云边协同等。其典型的业务特征为~100G级带宽、~1ms以下时延(1ms时延圈)、高可靠等。这些场景主要适合于中短程传送的城域、本地、一体化片区等。

此外,通算类型的灾备、多云互联等业务场景在不同的传输距离要求下有不同的业务特征和需求,需视具体情况匹配相应的网络能力。

### 1.1.3 DCA高品质入算场景

入算是需求和业务特征最为多样化的算力场景大类。它涵盖了普遍公众2C用户、一般政企类用户、垂直类工业2B用户、超大2G/2B用户、特型/科研用户等众多面貌的用户类群,业务节点呈现广覆盖特征,数据传送距离涵盖短、中、长等所有传输距离。

a) 智算类型主要包含训练类的样本面拉远训练、大样本传送等场景(中短距的同城异地样本面拉远以及中长距的“东数西训”、数据快递),推理类的用户侧边缘入算、企业边缘入算或一跳核心入算等以及车联数据算训等。

b) 通算类型主要包含各类公有云、私有云、混合云的用户就近入算以及远距的云存、东数西存等。

c) 超算类型主要面向各类科技和知识创新诉求(如基础科学、新兴技术、新质生产力工业服务业等)对超算的调用和应用等场景。

## 1.2 业务核心要素变化趋势

以算力为代表的新兴业态驱动着业务核心要素发生显著、广泛且持续的变化,这些变化主要有流向流量、节点分布、带宽体量和性能要求等几个方面。

### 1.2.1 业务流向流量演变趋势

业务网络模型由“核—网—端”(业务核心—承载网络—接入网及用户端)演变为“算/云—网—边—端”(核心算力/云池—承载网络—边缘算力/云池—接入网及用户端),驱动了业务流量逐步由传统的南北向

的流向模型演变为东西向、分布式、扁平化的弹性流向模型。业务流量也由此从逐级收敛、易于预测的模式转变为锯齿型的、突发性的流量模式,连接时长存在不确定性,场景差异度大,流量变得不易预测。

### 1.2.2 业务节点分布趋势

在算力业务崛起的同时,叠加城市集群化的战略规划(如长三角、珠三角、京津冀等一体化发展战略)以及算力高耗电特性与城市用电指标、节能减排压力之间的突出矛盾等因素,节点分布形态将不断地由传统的以业务流和人流汇聚方向为重、业务核心围绕着人文地理重心部署的“聚中式”的形态,转变为新兴的核心算力重点部署在地理外围(城市外围、国家西部区域等),同时构建下沉式、分布式、扁平化的边缘算力池或集群,形成“外展式”的节点分布特征,并通过网络实现算力和业务的疏导。

### 1.2.3 网络带宽和性能要求的发展趋势

与AI大模型参量规模不断快速提升形成鲜明对比的是,一方面,GPU集群在单体AIDC内受到供电、制冷模式、机位空间等设施的制约越发明显,在迈向万卡、超万卡时代时,势必转向以分布式算力集群训练为基础;另一方面,大量分散的中小规模AIDC算力节点也存在未全时满负荷运行的情况,这造成了一定的资源浪费,有必要通过网络构建异质融合的分布式算力训练池。

这2种场景都会带来海量的DCI流量,并对传输性能提出极高的要求。如何满足广域RDMA对underlay网络提出的高带宽、无损性(低时延、零丢包、高可靠等)等性能要求,支撑算力参数面和数据面在跨域间的高效传输需求,实现“以网强算”,成为提升大模型训练效率的关键因素。

a) 高带宽:节点间算力流量主要考虑Scale out互联层面的需求,以数据并行和流水线并行流量为主。在无收敛时,参数面网络单链路带宽至少需达到200~400Gbit/s,数据面网络单链需100Gbit/s以上带宽。若考虑10倍收敛比且不做其他架构优化,2个广域的千卡集群参数面互联流量即可达到百Tbit/s级。

b) 无损性:在规模给定(GPU单卡算力×总卡数不变)且带宽确保的情况下,集群的有效算力将受到线性加速比和有效运行时长的制约,对于分布式算力集群而言,两者都受到网络性能显著且直接的影响,其直观表现为吞吐率的高低<sup>[1-2]</sup>,而本质因素则是网络(含RDMA网络及承载网络)的时延和丢包率等因素

(如单流传输时,时延由 1 ms 增加到 10 ms,吞吐率下降约 90%;2%的丢包将使 RDMA 吞吐率下降为 0)。

对于因端侧数据发送带宽不均衡(典型情况如 ECMP 多路径时流量哈希冲突等)等原因造成的链路拥塞、出端口拥堵等问题,应主要考虑在 Overlay 网络(广域 RDMA、无损网络等)层面进行优化解决;而链路上的确定性可靠和无损/低损的传输承载性能应由 Underlay 承载网络来保障。

## 2 承载网络演进策略分析

业务核心要素的变化将驱动承载网络在拓扑结构、组网灵活性、容量能力、无损性能、智能调度能力等各方面发生深刻的变革。

### 2.1 网络拓扑演进策略

网络拓扑结构演进的驱动力主要来自扁平化的业务流向变化和扩展式的节点分布变化。主要诉求包括提升算力节点间一跳直达能力以及客户侧至算力侧的一跳入算能力;降低路由迂回以及跳接数;降低网络时延以改善或控制广域 RDMA 的吞吐率性能;提升网络多路可选性以满足保护需求,支撑多因子算路的底层能力等。具体的措施主要可考虑以下 2 点。

a) 优化网络架构:采用更为扁平化的网络拓扑结构,进一步减少网络层次,以降低网络延迟,提高带宽利用率,实现更为高效的数据传输。

b) 提升网络的 Mesh 化程度:构建网络各节点间的直达路径,以提高网络的可靠性和灵活性。Mesh 化网络可以提供多条冗余路径,减少单点故障,并支持更灵活的流量调度。

### 2.2 组网灵活性演进策略

组网灵活性演进的驱动力主要来自算力应用场景的多样化以及算力节点和流量的泛在化、动态涨落、不确定性高等特点。承载网在保持广覆盖优点的同时,需要具备更强的组网灵活性与可扩展性,以适应不同规模和类型的算力业务的灵活接入和承载需求。

在这一方面,相比 IP 网络的弹性组网能力,基于光传送网构建的全光底座面临着更为严峻的挑战。作为综合性基础承载网,它一方面需要承接泛在广覆盖的、异类异质的业务系统和节点以及其拓扑连接和颗粒、流量、时延、SLA 策略等差异化极大的承载需求;同时,算力等新兴业态正驱动着业务核心要素快速且动态地演进。另一方面,它受到光网底座的技术基础

(即 ROADM/OXC 维度能力)的约束。

因此,应不遗余力地加大光传送网的高维/超高维 OXC、云池化波分等技术的研发力度,扩大其应用范围,以满足核心节点间 Full Mesh 互联一跳直达,泛在广覆盖业务节点便捷接入等方面的灵活组网需求。

### 2.3 网络容量演进策略

网络容量能力演进的驱动力主要来自算力链路颗粒和业务体量的发展趋势。承载网需要扩大 400G 波分的部署规模,加快应用进程,提升网内 400GE 客户侧接口占比。当前 400G 波分主流技术对比如表 1 所示。

与 400G 波分应用部署策略密切相关的另一个关键因素是 C+L 一体化的成熟度。C+L 一体化分为 WSS 一体化、OTU 一体化和 OA 一体化。这 3 者的特征、价值和成熟度(预测)情况如下。

a) WSS 一体化:单块 WSS 板卡支持 C+L 频段。可减少光层器件、连纤数量,实现 C+L 波道整体交叉调度。目前该技术已成熟。

b) OTU 一体化:单块 OTU 板卡能够实现全 C+L 波长可调,这极大地方便了 WSON 运维工作,可通过改变频率的方式规避波长冲突。预计该技术还需 1~2 年才能成熟。

c) OA 一体化:单块光放板卡可放大 C+L 全频波道。可减少光层器件、连纤数量,实现 C+L 波道整体交叉调度。预计该技术再需 3 年以上才能成熟。

近期部署 C+L 一体化的 400G 系统在工程建设和运维方面将会面临以下 2 个难点。

a) SRS 效应导致的功率转移不可忽视,需采用填充波等措施(但这会使运维复杂度有所上升)。

表 1 400G 波分主流码型简要对比

400G OTN 主流码型指标对比		400G 16QAM	400G PCS-16QAM	400G QPSK
信号调制载波	波特率	~60+G	~90+G	~130+G
容量	波道间隔	75 GHz	100 GHz	150 GHz
	C 波段波道数	80	60	40
	C 波段+L 波段波道数	80+/ ("/"表示结果待定)	60+60	40+40
场景	最大单纤容量	32T@C	24T@C, 48T@C+L	16T@C, 32T@C+L
	工程距离 (含 ROADM 跳站)	~300 km	~500 km	~1 000 km
	典型应用范围	城域网/本地网	城域网/本地网/二干	一/二干

b) 一体化 OTU 尚未成熟,若启用 WSON 功能,其倒换恢复的难度上升;需在启用 WSON 功能的传输系统上合理预规划波道(这会使运维复杂度有所上升),或等待该技术成熟。

不同的场景会产生不同的约束条件,进而导致不同的技术选型策略或倾向。综合以上因素,近期 400G 技术选型的简要建议如下。

a) 在城域/本地网内,无需以远传输距离(>~500 km)作为一级约束条件。因此,可追求更高的调制阶数,从而在单波速率给定的情况下可放松对更高波特率的刚性要求,进而获得更优的频谱效率。基于此,可充分立足于 16QAM/PCS 码型的频谱效率优势,发挥 C 波段快速部署的特性。

b) 一、二千系统,尤其是长距段落中,需重点考虑无电中继传输距离的约束条件。在此条件下,为满足系统 OSNR 容限,需在调制阶数上适当进行妥协。进而,为了获得同等的单波速率,需进一步增大波特率,这会直接导致波道间隔的提升,最终不得不以牺牲一定频谱效率为代价。因此,可选用 QPSK 码型,以发挥其长途传输特性,并在 C+L 波段下进行部署;近期可基于 C/L 波段隔离策略合理规划波道,以支撑 WSON 智能调度。

与此同时,应推动 800G 及以上速率的波分系统、空芯光纤等新技术新材料的研发工作,并对其部署应用策略进行研究,进一步提升高速海量的承载能力。

## 2.4 网络无损性能演进策略

网络无损性能演进的驱动力主要来自算力广域 RDMA 无损网络对 Underlay 网络的确定性传输性能诉求。

a) 光传送网的基础光域技术(如 OXC/WSS-ROADM、波分等)具备持续稳定的低时延(如微秒~十微秒级的设备处理时延、等价于裸纤的路由传递时延等)、超低丢包率(纠前 $\sim 10^{-12}$ )、高可靠性(具备 OLP、OMSP 等多种 50 ms 以内系统级的保护倒换策略,WSON 等秒级的波长路由动态调度策略以及刚性的业务管道隔离/网络虚拟切片技术等)。

为进一步发挥光传送网的无损传输性能优势,一方面可采用 Mesh 化的 ROADM/OXC 组网方式,进一步提高网络的扁平化程度(实现业务直达、降低时延)和多路由能力(抗链路故障能力);另一方面应加大对空芯光纤(HCF)的研发投入。相比普通光纤,HCF 中光的传播速率可提升约 1.5 倍,能够发挥显著的超低时

延性能优势。

b) 相比而言,IP 承载网在此方面面临的挑战更为严峻,需要进一步深入研究“确定性网络”的相关技术和应用策略。一方面,通过严格控制传输路径和时间参数,确保数据传输时延的可预测性;另一方面,采用智能流量管理机制,有效预防和处理网络拥塞,确保数据稳定传输。

## 2.5 智能调度能力演进策略

在以算力为网络核心的背景下,为获取能效、性能和利用率这 3 个制约因素的动态较优平衡点,需要树立“算网一体”的系统性思维观念。在算网融合的目标指引下,构建算力和网络有机统一的资源编排体系,构建实现算力通过网络平滑迁移、扩展和调度的协议机制。

算力网络的调度技术旨在合理分配和高效利用算网资源,提高资源利用率、减少不合理的资源闲置或浪费。例如,将多个分散的算力资源池进行整合并统一使用,以提升算力规模,扩大使用范围,进而保障任务的高效执行。

## 3 算力承载网络架构研究

### 3.1 总体目标及架构模型

算力承载网络的总体目标是构建智能敏捷、按需随选的算力承载架构。

算力业务除“单节点/单局址”的算内场景以外,在其他诸如入算、算间、分布式算内等涉及远程算力业务承载、调度的场景下,均需依赖传输承载网提供高效、可靠、可确定、敏捷智能的连接能力。

面向算力的传输承载网,需将传统的光传送网和 IP 承载网进行有机整合,演进为全光网底座+灵活承载层+智能路由层的融合架构,进而构建全新的广域算力体系下的 Underlay 网络。网络演进目标如图 1 所示。

### 3.2 架构组成及关键技术

#### 3.2.1 全光网底座

全光网底座主要基于光传送网的高维光层组网、高速传输等技术进行构建,具备扁平、灵活、高速海量、高效绿碳、无损可靠健壮的特征。

a) 高维光层组网技术:基于 32 维及以上的 OXC 来构建站间一跳/短跳直达、低时延、灵活组网、无阻塞交换的 ROADM 网络,并可灵活配合 WSON、OMSP、OLP 等系统级保护倒换恢复机制,提供高安全可靠的

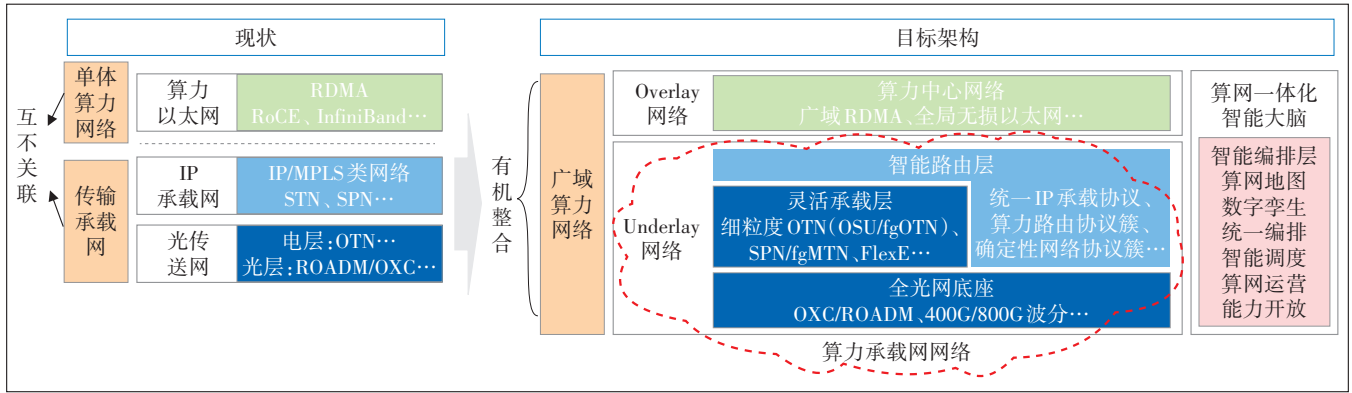


图1 面向算力的承载网络演进目标及架构

网络保护能力。

b) 高速传输技术: 基于400 Gbit/s、800 Gbit/s 或更高速率的高速/超高速单波技术、C+L等频谱扩展技术、新型光纤技术等, 构建具备高速连接、海量承载、绿色低碳的传输通道资源池。

全光网底座的技术应用对网络特性和内涵的映射关系如图2所示。

### 3.2.2 灵活承载层

面向多业务综合承载需求, 主要基于光传送网电

层/IP承载网的新型切片技术, 构建具备敏捷接入、智能适配、灵活综合承载能力的中小颗粒业务通道。其主要技术路线如下。

a) 由OTN发展而来的细粒度OTN技术(如OSU、fgOTN等)以及由SPN发展而来的fgMTN技术。细粒度OTN和fgMTN等技术继承、延续并提升了原有的分组增强型OTN或SPN的中小颗粒灵活承载能力, 可高质量承载TDM、分组业务, 增强业务调度灵活性。在服务于算力网络时, 可满足10M~10G级别中小颗粒

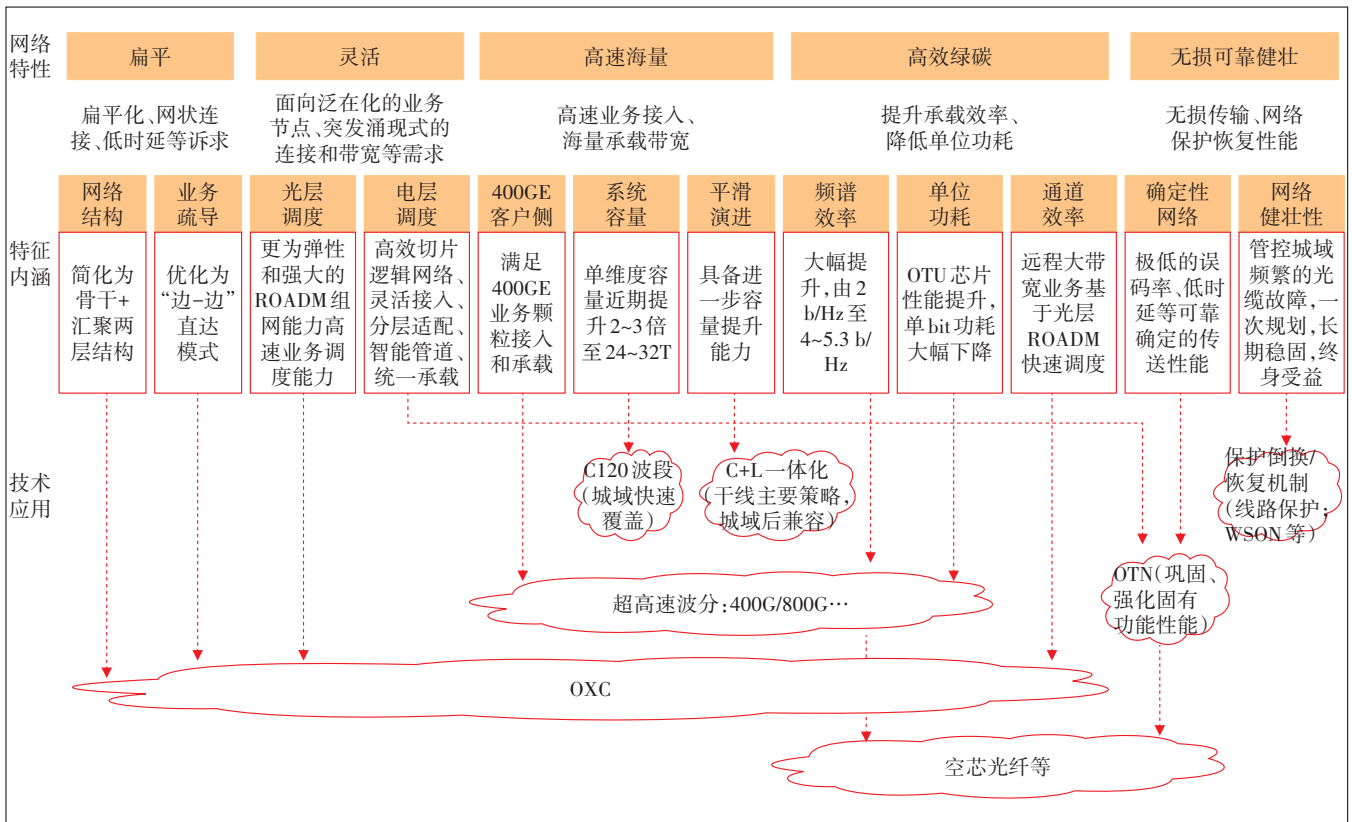


图2 全光网底座网络特性及技术应用映射关系示意

客户的信号高效承载需求<sup>[1,3]</sup>。

b) 由传统以太网发展而来的 FlexE 技术等。FlexE 通过 TDM 思路实现业务速率与物理通道速率解耦<sup>[2]</sup>, 在提供比传统以太网更加灵活的带宽颗粒度的同时, 还具备使物理通道向更高速率演进的潜力。它基于专用的物理通道, 且同业务需求特性相适配, 可实现确定性带宽、带宽按需分配等性能和功能。

c) 基于 SRv6 的 IP 网络层次化切片能力等, 识别精细的业务特征(如差分服务代码点、应用响应网络标识等), 并据此进行分流、引入隧道操作。在资源保障、转发标识、路径规划等层面上综合进行网络切片层次化, 在满足统计复用需求的基础上实现资源共享能力<sup>[1,4]</sup>。

通过灵活承载层, 系统可针对不同 SLA 业务特性需求的算力业务智能地创建对应的虚拟切片网络, 既可组建刚性、硬隔离的传送管道, 又可组建弹性、统计复用、低成本的传送管道。

### 3.2.3 智能路由层

智能路由层主要针对不同特性、动态变化的复杂算力业务对网络资源实时调用或切换的需求, 重点解决业务感知、信息通告、智能路由、确定性网络性能等方面的问题<sup>[1-4]</sup>。它是将算力和网络紧密结合、实现算网一体的重要功能层级。其主要技术路线如下。

a) 算力路由协议簇: 立足于解决算力度量、信息通告、多因子算路等问题。通过将算力信息引入网络路由域, 同时考虑网络与计算资源状态的实时匹配情况, 实现两者的联合优化, 从而满足时延敏感和计算敏感业务的需求。此外, 它还具备上层业务感知和网络资源可编排能力, 支持灵活、可编程的数据面转发以及通过确定性协议机制(如 CSQF 等)确保的传送转发性能等<sup>[2-4]</sup>。

b) 确定性网络协议簇: 立足于解决 IP 网络性能质量问题, 包括 DetNet、DIP、TSN 以及衍生自 SRv6 的增强的确定性传送转发或统一 IP 承载协议等。它能够显著提升 IP 网络在丢包率控制、时延、抖动等方面的表现, 并增强队列调度、时隙映射、流量整形、资源预分配等方面的能力<sup>[1-4]</sup>。

智能路由层使承载网络获得了对算力业务的感知能力。该层可与 Overlay 网络协同形成流控机制, 以解决广域场景下吞吐量下降和拥塞等问题; 同时具备确定性转发能力, 以保障算力性能指标; 此外通过光+IP 协同控制机制统一调度 Underlay 网络资源, 实现性

能可靠确定、调度智能灵活的广域无损传输。

### 3.2.4 智能编排层

通过智能编排层来实现算网一体的智能化资源编排、联动管控和协同调度, 使承载网络具备统一的需求-性能匹配、全系统资源编排和整合、跨域跨层协调调度等机制和能力。基于此, 承载网络能够进一步动态感知和协商算力业务的特性要求和资源需求(带宽、时延、链路拓扑/路由连接、SLA 等级、时序/时长等)等参数, 将全光网底座的广泛覆盖、全光调度和扁平高效组网能力, 灵活承载层的敏捷综合承载、平滑扩展和灵活拆建/复用能力, 智能路由层的算力通告、性能确定、网络调度、资源编排等能力相结合, 构建高度弹性敏捷、按需随选的 Underlay 网络资源池。

## 4 结束语

算力承载网络是当前算网融合趋势下学科交叉创新的热门领域, 有着广阔的应用前景和发展潜力。基于高效的算力承载网络, 可实现分布式的智算中心聚合, 形成虚拟融合的超级 AIDC 算力节点, 也可实现分散异构、碎片化的小型或边缘算力节点汇聚成池, 有效支撑 AI 大模型对大规模算力资源的指数级增长的需求, 将点状的城域算力节点构建成算/云、网、边、端协同的层次化算网合一体系。

### 参考文献:

- [1] 中国移动通信集团有限公司.“九州”算力互联网(MATRIXES)目标架构白皮书[R/OL]. [2025-05-20]. <https://baijiahao.baidu.com/s?id=1797489266186712566&wfr=spider&for=pc>.
- [2] 中国电信股份有限公司研究院.云网一体信息基础设施白皮书[R/OL]. [2025-05-20]. <https://www.sgpjbg.com/baogao/602970.html>.
- [3] 中国联通研究院.算力时代的全光底座白皮书[R/OL]. [2025-05-20]. <https://baijiahao.baidu.com/s?id=1769035102862374778&wfr=spider&for=pc>.
- [4] 中国移动研究院.算力感知网络(CAN)技术白皮书[R/OL]. [2025-05-20]. <https://baijiahao.baidu.com/s?id=1769558828959350396&wfr=spider&for=pc>.

#### 作者简介:

张纬卿, 高级工程师, 长期从事光传输网络、数据通信的咨询、规划与设计工作。

