ICS 33.040.40
CCS L78

T/NIDA

全 球 固 定 网 络 创 新 联 盟

T/NIDA-005-2025

面向算力业务的城域网络建设技术要求

Technical Requirements for Metropolitan Area Network Construction Oriented to Computing Power Services

2025-09-26 发布 2025-09-26 施行

全球固定网络创新联盟（NIDA）发布

目　录

前　言

本文件按照 GB/T 1.1-2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利权和著作权。本文件的发布机构不承担识别专利和著作权的责任。全球固定网络创新联盟不对标准涉及专利的真实性、有效性和范围持有任何立场；不涉足评估专利对标准的相关性或必要性；不参与解决有关标准中所涉及专利的使用许可纠纷等。

本文件由全球固定网络创新联盟技术委员会提出并归口。

本文件起草单位：中国电信股份有限公司研究院、中关村超互联新基建产业创新联盟、华为技术有限公司、中兴通讯股份有限公司

本文件主要起草人：朱永庆 袁博 胡泽华 董杰 朱海东 吉晓威

面向算力业务的城域网络建设技术要求

1 范围

本文件规定了全球固定网络创新联盟中面向算力业务的城域网络建设技术要求。

本文件适用于面向算力业务的城域网络建设，主要应用于指导城域网络规划、设计和验收。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

3 术语和定义

3.1 算力指标定义

3.1.1 有效算力率（Effective Computational Power Rate）

本文中使用“有效算力率”泛指智算集群中衡量数据中心或计算平台实际计算能力利用效率的指标。它通常指的是实际可用于计算任务的算力与总可用算力的比值。

3.1.2 算力可用率（Computing Power Availability Rate）

本文中使用“算力可用率”泛指在训练任务过程中，计算资源实际可用于计算任务的时间与总时间的比率。它反映了计算资源的可用性，是衡量数据中心或计算平台资源利用率的一个重要指标。

3.1.3 集合通信效率（Collective Communication Efficiency）

本文中使用“集合通信效率”泛指在并行计算环境中，多个计算节点（或进程）共同参与的通信操作的效率。这些通信操作包括Broadcast、Gather、All-Gather、Scatter、Reduce、

All-Reduce和All-to-All等。集合通信效率直接影响到并行计算任务的性能。

3.2 网络术语定义

3.2.1 网络通信效率（Network Communication Efficiency）

本文中使用“网络通信效率”泛指智算数据中心网络环境中数据传输的效能，通常用来衡量数据在网络中传输时的速率、准确性和资源利用情况，包含吞吐率、丢包率、时延、带宽利用率等多个方面。

3.2.2 吞吐量（Throughput）

网络在单位时间内成功传输的数据量，通常以每秒比特数（bps）来衡量。

3.2.3 时延（Latency）

本文中时延指针对不同计算任务网络需要保障的端到端时延。

3.2.4 丢包率（Packet Loss Ratio）

未发送成功报文个数占总报文个数的比例。

3.2.5 带宽利用率（Bandwidth Utilization）

网络带宽被有效使用的比例，高带宽利用率意味着网络资源得到了充分的利用。

4 缩略语

下列缩略语适用于本文件：

AI：人工智能（Artificial Intelligence）

AIDC：面向智算的新一代数据中心（Artificial Intelligence Data Center）

ASIC：专用集成电路（Application-Specific Integrated Circuit）

CaaS：计算即服务（Computing as a Service）

CPU：中央处理器（Central Processing Unit）

DC：数据中心（Data Center）

ECN：显式拥塞通告（Explicit Congestion Notification）

EDN：增强确定性网络（Enhanced Deterministic Network）

FLOPS：每秒浮点运算次数（FLoating-point Operations Per Second）

FPGA：可编程逻辑门阵列（Field Programmable Gate Array）

GPU：图形处理器（Graphic Processing Unit）

MTTR：平均故障修复时间（Mean Time To Repair）

NPU：神经网络处理器（Neural network Processing Unit）

PUE：电能使用效率（Power Usage Effectiveness）

PFC：基于优先级的流量控制（Priority-based Flow Control）

POD：分发点（Point Of Delivery）

POP：接入点（Point Of Presence）

QoS：服务质量（Quality of Service）

RDMA：远程直接存储器访问（Remote Direct Memory Access）

SLA：服务水平协议（Service Level Agreement）

TB：太字节（TeraByte）

ZB：泽字节（Zettabyte）

5 面向算力业务的城域网络概述

5.1 算力产业发展趋势

随着通算、智算、超算技术的快速发展和广泛应用，算力需求呈现爆炸式增长。根据《IDC中国加速计算服务器半年度市场跟踪报告》分析，2025 年中国智能算力规模预计将达到 1037 EFLOPS，2028 年将达到 2782 EFLOPS，年均复合增长率达到 46.2%。依托庞大市场优势，我国算力水平和供给能力显著提升，形成了体系完整、规模庞大的产业体系。当前，算力产业呈现出多元化、集约化、智能化和普惠化的发展趋势，成为支撑千行百业数字化转型的关键引擎。

— 算力以往主要依赖 CPU 实现通用计算，广泛应用于办公、数据库、基础业务等场景。随着人工智能，尤其是深度学习和大模型技术的兴起，传统 CPU 在高并发、低延迟的矩阵运算方面已难以满足性能需求。GPU 凭借其强大的并行处理能力，迅速成为AI训练和推理的主流算力载体。与此同时，NPU、FPGA、ASIC 等专用芯片不断涌现，

针对特定 AI 任务实现定制化加速和优化。此外，超算算力在气象预测、生物医药、航空航天等关键领域仍持续发挥关键作用。未来，多种算力形态将协同共存，逐步形成多层次、异构融合的算力体系。

— 面对算力需求的快速增长，传统“烟囱式”数据中心建设模式已难以为继，资源分散、利用率低、区域性失衡等问题制约了算力效能的有效发挥。为此，国家大力推进“东数西算”工程，通过构建八大国家算力枢纽和十大国家数据中心集群，推动算力资源在全国范围内的统筹布局和协同调度。运营商和云服务商也在积极构建“算力专网”、“算力交易平台”等新型服务模式，实现算力资源的可视化、可调度、可计量，打造“按需分配、随用随取”的算力服务模式。通过上述举措，算力正从基础设施资源向综合服务能力转变，进一步加速产业的集约化演进。

— 随着算力系统规模的持续扩大，传统人工运维已难以应对日益复杂的资源调度与故障处理。智能化成为提升算力运营效率的关键路径，AI 技术正逐步应用于算力资源的全生命周期管理：在调度层面，利用强化学习和预测模型，实现任务的智能分发与负载均衡，避免局部过载或资源闲置；在运维层面，基于大数据分析和异常检测算法，实现故障预警、根因分析和自动恢复，显著提升系统可用性；在能效管理方面，通过 AI 动态调节服务器频率、冷却系统功率和供电策略，有效降低 PUE，推动算力基础设施的绿色低碳、可持续发展。

— CaaS 模式日益成熟，算力正在从“自建自用”向“服务化、平台化、市场化”加速演进。对于多数企业和研究机构而言，自行建设和维护 AI 基础设施需要投入大量的人力和物力。在此背景下，算力租赁凭借即插即用、弹性可扩展的算力供给方式，显著降低了企业的算力获取难度和使用成本。企业可通过租赁第三方 AIDC 中的 AI 基础设施，更快捷地获取所需算力资源，从而加速技术研发和创新。算力租赁有效降低了企业的算力使用门槛，推动算力技术的普惠化发展，为算力服务的广泛应用和持续创新提供了强有力的支撑。

5.2 算力业务关键场景

算力业务是指通过提供、调度和管理算力，满足个人、企业或机构对高性能计算需求的服务。伴随人工智能技术的快速发展，大模型训练和推理成本快速下降，驱动算力需求快速增长，海量数据入算、存算分离拉远训练、跨集群协同训练、云边协同训推等算力业务不断

涌现。在此背景下，城域网络作为连接用户与算力资源的关键基础设施，为各种算力业务提供了关键网络支撑，确保使用异地算力资源的企业可以获得接近本地部署的算力使用体验。

5.2.1 海量数据入算

随着 AI/HPC 的迅猛发展，数据规模正在以前所未有的速度增长，企业单次向算力中心传送的数据集可达数百 TB。IDC 预计，2025 年全球将产生 213.56 ZB数据，到 2029 年将达到 527.47 ZB; 其中, 中国市场 2025年将产生 51.78 ZB 数据, 到 2029 年增长至 136.12 ZB。当前，众多企业仍依靠邮寄硬盘的方式进行大规模数据的搬运，每年都有 PB 级数据通过硬盘搬运/邮寄方式传送到 AIDC 进行模型训练。这种模式不仅效率低，还面临着硬盘损坏与数据丢失的风险。而基于网络传送的方案仍存在不足，百兆专线耗时长，而万兆专线/OTN 专线成本高，亟需对网络进行升级，提供更为高效且具性价比的数据入算服务。

5.2.2 存算分离拉远训练

数据安全要求广泛存在于多个领域的智算场景中。如汽车制造业涉及的碰撞实验和事故数据，政务领域涉及的官方文件、公民身份信息及法人资料等敏感信息，这些数据均具有较高的安全标准。这些企事业单位对样本数据有严格的安全标准，明确要求核心数据存储在其所在园区或单位内。在租赁第三方算力资源时，这些企事业单位在坚持数据本地化存储原则的同时，还需要确保数据在模型训练过程中不被泄露。因此，在该场景下，算力资源节点与数据存储节点需要跨网络部署（通过网络实现两者之间的高频实时交互），仅在模型训练时分批上传所需的样本数据，样本数据用完即弃。

5.2.3 跨集群协同训练

大模型 Scaling Law 持续生效，过去十年间大模型的算力需求增长约 100 万倍，预计未来仍将保持每年 4 倍以上的增长。考虑到单个数据中心的算力规模受电力供应、机房空间等多重因素的制约，为满足大模型快速增长的算力需求，需要推动多 AIDC 跨网络协同训练，整合分布在不同地理位置的分散算力资源。同时，我国智算中心规模普遍偏小（规模为 100-300 PFLOPS 的小型智算中心占比超 70%），并且往往分散在不同的数据中心、科研机构、地方政府和云服务商。因此，基于网络整合零散的社会算力，也有助于打破地域、机房资源、服务商等限制，构建统一、高效的算力服务平台。

5.2.4 云边协同训推

大模型训练与推理成本的显著降低，带动企业通过本地部署训推一体机实现大模型的快速落地应用。然而，企业本地算力池面临扩容难、维护成本高等问题，难以满足大模型微调和推理不断增长的算力需求。通过企业本地算力与云端租赁算力之间的高效协同，满足企业灵活扩展算力资源的需求，成为更高效、便捷且兼具性价比的方案。云边协同训推方案基于Split Learning 部署模式，将模型切分到本地和云端算力资源池中并行处理，由网络实现梯度、KVcache等数据的高效同步。此外，该方案可结合输入、输出层的本地化部署，保证样本数据不出园区，满足金融、医疗等数据敏感客户的数据安全要求。

5.3 城域网络面临挑战

城域网络连接了区域内的异构算力资源和多样化用户终端，是支撑人工智能产业持续发展的关键基础设施。城域网络不仅服务于客户分散部署的园区与数据中心之间的互联，同时也为基于算力租赁的算力服务提供基础网络支撑。为构建类似城市水、电的算力服务，实现“一点接入、即取即用”，城域网络面临以下三个核心挑战：

- **挑战1：数据流通的高效性与稳定性瓶颈：**AI 大模型训练与知识库构建通常需要TB/PB 级数据，这对城域网络的吞吐量提出了更高的要求。同时，大模型参数同步过程对网络性能要求极高，而低质量的网络可能导致数据传输带宽不足、时延过大或者频繁丢包等问题，从而影响计算资源的可用性。此外，MCP、A2A 等新型计算模式的快速发展，要求城域网络能够高效疏导智能体之间快速增长的东西向流量，并能够保障智能体之间信息交互的高可靠性。
- **挑战2：运营维护体系难以适配 AI 业务动态特性：**AI 业务流量呈现“大象流与小微流混合”、“突发与常态交织”特征，传统固定带宽机制易导致网络拥塞或资源闲置。同时，AI 业务对可靠性容忍度极低，微小故障就可能导致任务重置，直接影响用户体验。传统依赖人工排查的运维模式，MTTR 通常在分钟级，难以满足 AI 业务对高时效性的要求。此外，现有运维体系在算力调度能力方面存在明显短板，无法实现跨节点、异构的算力资源统一感知与弹性调度，导致算力资源利用率偏低。
- **挑战3：数据安全与可信机制构建难度高：**AI 业务涉及大量敏感数据，传统基于流量过滤的 AAA 体系和基础加密技术难以防范定向攻击，难以保障数据全生命周期安全。

城域网络同时承载 ToH、ToC、ToB 业务，需要在保障业务隔离性的同时，灵活适配不同业务的性能需求。而传统 VLAN 隔离或 VPN 技术，难以动态调整业务的带宽、路径等网络参数，资源共享与隔离的平衡成为难题。此外，区块链、量子加密等新兴安全技术尚未与城域网络实现有效适配与融合，导致可信架构构建成本高、周期长。

因此，面向算力业务的城域网络（以下简称算力城域网）建网需要在通信效率、可维护性、可靠性、安全性等方面来保障有效算力率、算力可用率、集合通信效率等关键算力指标。

6 面向算力业务的城域网络架构

6.1 整体架构

算力城域网应基于积木式架构实现模块化组网，支持网络的按需灵活调整与异构组件的兼容接入，保障网络的平滑升级与可持续演进。当业务需求较少或城市规模较小时，可简化算力 POD、算力 POP、算力互联区等组件，基于 Spine-Leaf 等网络架构进行城域网络改造与建设，以快速满足初期业务承载需求。

根据功能的不同，算力城域网可分为算力 POD、算力 POP、算力互联区三大模块化组件以及运营管理系统。总体网络架构如图 1 所示：

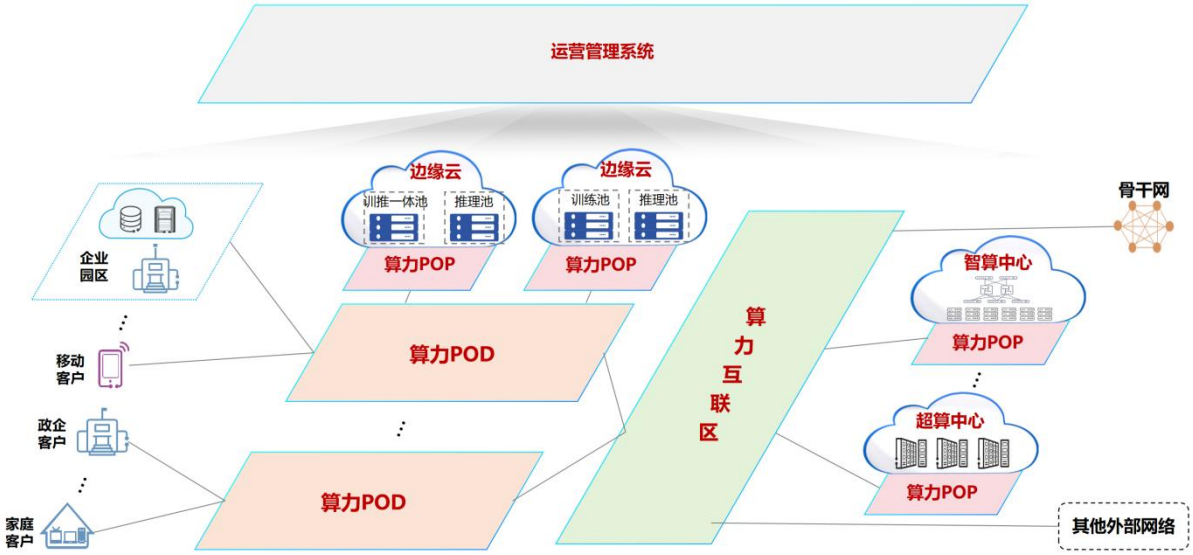

图 1 算力城域网总体架构

— 算力 POD：作为城域范围内全业务融合承载的核心单元，实现各类算力用户的广泛覆盖和统一接入，以及深/浅边缘算力池的快速接入。

— 算力 POP：联接城域网络与异构算力资源池，实现二者在参数面、样本面、业务面网络的标准化、快速对接。

— 算力互联区：负责城域网络与异构算力资源池及骨干网、核心网等外部网络的互联互通，统一疏导算力 POD、算力资源池之间的内部流量，以及城域网进出口流量。

— 运营管理系统：作为城域网络的“调度大脑”与“管理中枢”，实现算网资源的统一管控、协同调度及智能运维。

通过三大模块化组件之间的高效协同，确保算力业务在城域内的高效承载。算力 POD 作为用户接入入口，通过与算力 POP 的高速互连，构建用户至算力资源池间的高效传输通道；算力 POD 与出口功能区联动，实现用户与跨 POD、跨域算力资源池的端到端无损连接。与此同时，三大组件通过标准化接口协议与运营管理系统无缝协同，形成“资源可视、业务可管、连接可控”的统一管控体系，支撑算力资源的动态感知、智能编排与全生命周期运维管理。

6.2 各模块核心能力

6.2.1 算力 POD

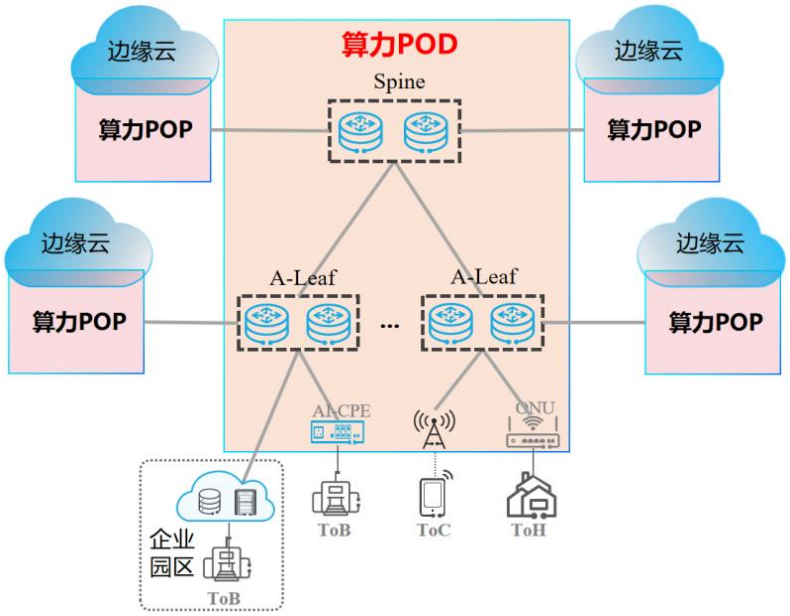

图 2 算力 POD （以 Spine-Leaf 组网为例）

算力 POD 是算力城域网的边缘接入层，负责客户终端、企业分支及家庭用户的融合接入。通过设备逐级收敛汇聚流量，形成覆盖广、弹性灵活的算力服务入口。其核心功能包括：

— 可根据用户位置、行政区域、AIDC 服务范围等因素灵活部署，支持算力用户终端、企业分支站点的泛在接入，并涵盖光纤、PON、5G 等多种接入介质。

— 可实现固、移、云、算业务的统一接入、融合承载，保障各类业务流量的无阻塞、快速转发。

— 可为用户提供高吞吐、低延迟、低丢包、高可靠的算力通道，确保数据传输质量。

— 可结合算力 POP 快速接入边缘算力资源池，实现边缘算力的池化利用和灵活调度，为客户提供低时延、高体验的算力服务。

6.2.2 算力 POP

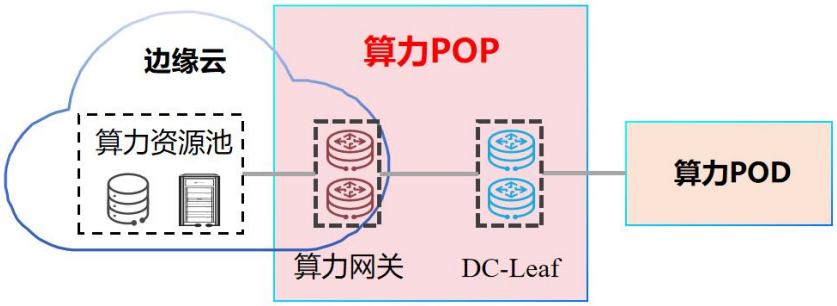

图 3 算力 POP

算力 POP 通过专用 Leaf 设备与算力网关之间的标准化组网，实现算力城域网与算力资源池的快速对接，支撑多业务一体化承载与算力灵活调度。其核心功能包括：

— 可实现城域网络与自有、第三方异构算力资源池的样本面网络、参数面网络、业务面网络之间的标准化、快速对接，保障“分钟级”算网业务开通以及“ ms 级”算网业务访问时延。

— 可基于 SRv6 技术实现入算、算间等流量的快速转发与业务端到端可编程，为用户提供高速、稳定的云网服务；

— 可与省级或区域算力 POP 互联，实现跨域算力资源池之间的高效连接，支撑跨域算力协同。

6.2.3 算力互联区

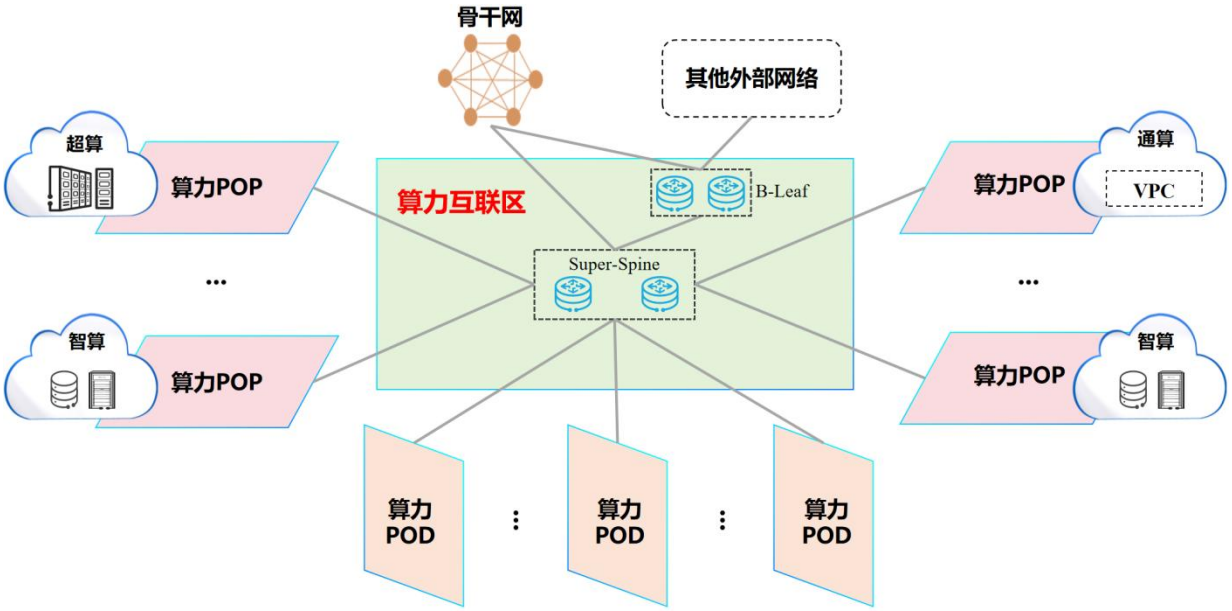

图 4 算力互联区 （以 Super-Spine、B-Leaf 组网为例。其中， Super-Spine 定位于承载多算力 POD 间、异构算力资源池间互联流量，以及算力 POD 与算力资源池间互访流量，同时与骨干网对接；B-Leaf 定位于承载跨域或者跨 POD 的算力业务流量，同时与其他外部网络对接）

算力互联区作为算力城域网的流量枢纽，实现算力 POD、算力 POP 与异构算力资源池以及骨干网、核心网等外部网络之间的互联互通。其核心功能包括：

— 可在不改变与骨干网/业务网等外部网络连接的条件下，实现算力 POD、算力 POP 等组件的灵活扩展。

— 可实现城域网络进出口流量、多算力 POD 间流量的统一、快速疏导，简化网络层级。

— 可实现异构算力资源池之间的高速互联互通，保障异构算力的高效整合与协同调度。

6.2.4 运营管理系统

图 5 运营管理系统

运营管理系统是统一调度、管理和优化网络资源与服务流程的综合平台，作为城域网络的“管控中枢”，包含资源调度、服务编排、性能监控、故障处理、运维管理等核心功能。对于算力城域网而言，其运营管理系统聚焦资源管理、算力调度、业务编排和运营保障四大核心功能：

— 资源管理：整合网络拓扑、带宽状态等网络信息与算力节点类型、负载状态等计算信息，实现算网资源的统一建模、动态感知与可视化管理。

— 资源调度：基于业务 SLA 需求与实时资源状态，智能分配最优算力节点和网络资源，支持任务迁移、弹性扩容与跨域协同，提升资源利用效率与服务质量。

— 业务编排：通过标准化模板与自动化流程，实现固、移、云、算多类型服务的快速部署与端到端贯通，提升业务上线效率与灵活性。

— 运营保障：结合数字孪生、AI 分析等智能运维手段，实现故障的预测、定位与自愈，以及安全风险的主动识别与快速响应，保障系统安全稳定运行。

7 面向算力业务的城域网络关键技术能力要求

为明确各模块需具备能力，本章针对算力 POD、算力 POP、算力互联区、运营管理系统的关键技术能力要求分别展开介绍。其中，为保障算力业务的端到端无损、确定性传输，算力 POD、算力 POP、算力互联区均需具备广域无损与确定性服务能力；运营管理系统需从资源分配、路径优化到 SLA 保障全方位保障算力资源利用效率与用户体验。

7.1 算力 POD 关键技术要求

算力 POD 可根据业务规模、AIDC 服务范围、行政区域等灵活设置，以适应业务快速增长需求。同时，算力 POD 可支持各类用户终端和企业分支站点的接入，实现固网、移动、云和算力业务的统一接入与融合承载，并保障区域内各类业务流量的无阻塞快速转发。为保障用户的高品质算力服务体验，算力 POD 还需提供低时延、弹性、无损的算力传输通道，实现算力的广泛覆盖、灵活调度与高效输送。

算力 POD 需满足如下关键能力要求：

a) 一线接入、融合承载：支持基于 SRv6/EVPN 统一协议栈，实现固网、移动、云服务及多元算力业务的一体化接入与承载。用户可通过单一接入点同时接入网络、云资源

及多个 AIDC，有效保障业务体验的一致性，并显著降低网络复杂度和运维成本，提升业务部署效率与用户体验。

(1) 支持专线、SD-WAN 等多种互联方式，并涵盖光纤、PON、5G 等多种接入介质，实现企业/家庭的本地局域网与第三方云/算力资源的无缝连接，构建逻辑统一、安全隔离的虚拟专网，实现数据与算力的高效、安全互通。

(2) 能够根据业务流量特征、算力和模型的需求类型、以及 QoS、安全等增值业务要求，自动识别并标记算力业务流。并通过与运营管理系统联动，为算力业务动态分配最优的转发路径，以及包括网络切片和带宽在内的各类网络资源，实现算力业务的高效承载与 SLA 保障。

b) 广域无损：具备广域无损传输能力，有效应对存算分离拉远训练、云边协同分布式训推等场景中 RDMA 传输对丢包的高敏感（0.1% 丢包即导致吞吐率下降 50%）。通过部署 400GE 高速端口、大缓存设备及精细化拥塞控制等技术，有效避免网络拥塞，抵御网络丢包，保障算力高效传输。

(1) 城域网设备应支持 400GE 高速接口，满足高吞吐、低时延的智算业务互联需求。同时，设备还应支持 GB 级大端口缓存和 K 级 RDMA 无损队列，并能够通过感知业务流量变化以优化缓存、队列分配策略，提升网络在高突发、高并发、高负载环境下的稳定性与传输效率。

(2) 算力 POD 应具备精细化拥塞控制能力，支持以业务/租户为单位动态分配设备缓存与优先级队列，并基于端到端/逐跳回溯方式向上游设备发送精准反压信号，从而实现细粒度的流量控制。同时，应支持 ECN、PFC 等拥塞控制机制以及相关优化，确保在高并发场景下实现多维度、细粒度的拥塞预防与快速响应。

c) 超高吞吐：具备超高吞吐数据传输能力，保障东西向/南北向流量的无阻塞疏导。同时，可基于报文深度识别（可识别到 IB 传输层信息）及 SRv6 流级调度技术，实现“大象流”自动识别、子流拆分与高效负载均衡，消除高带宽、长周期的大象流造成整体吞吐性能下降的风险，最大化链路利用率，保障网络吞吐率可达 90% 以上。

d) 高收敛比组网：具备高收敛比组网能力，以高效承载云边协同训推和零散算力整合业务。这类业务具备高并发、大突发流量等特征，对网络带宽提出极高要求。算力POD通过“智算缓存+智能调度”双重机制在实现高收敛比组网的同时，有效保障用户体

验：高速缓存模块可瞬时吸收流量峰值，缓解突发压力；同时，基于业务优先级动态调度，确保控制信令、任务数据等关键业务优先转发，避免计算资源闲置。通过按需引入 8:1、16:1、32:1 等高收敛比，大幅降低企业专线带宽需求，最终在实现带宽成本优化的同时，维持算效稳定，达成企业降本与用算效率的双重提升。

e) 确定性服务：具备确定性服务能力，通过 FlexE、IP 确定性切片等技术与 SRv6 可编程路径能力相结合，面向多租户提供软硬结合的层次化网络切片服务，为业务提供端到端差异化保障。同时，需支持“资源预留 + 路径编排 + 流量调度”多层次确定性能力，在复杂业务负载和大规模流量并发场景下，为算力业务提供确定性的低时延、低抖动和高可靠数据传输服务。

f) 安全加密：需支持 IPsec、Macsec 等多层次加密机制，确保用户数据在端到端传输过程中的机密性和抗窃听能力。同时，应支持国密算法和国际主流加密标准的高速加解密功能，保证在不降低转发性能的前提下实现全链路加密传输。此外，可结合量子加密以及可信执行环境等前沿安全技术，进一步增强网络的安全防护能力。

7.2 算力 POP 关键技术要求

算力 POP 随算力资源池建设，可实现算力城域网与异构算力资源池的参数面、样本面及业务面网络间的标准化、快速对接。同时，算力 POP 可支持端侧与网络侧的深度协同、RDMA 流量的无损传输以及多租户 SLA 的确定性保障，全面支撑存算分离拉远训练、云边协同训推等各种算力业务的稳定、高效运行。

算力 POP 需满足如下关键能力要求：

a) 标准化对接：通过 DC-Leaf 设备与算力网关的标准化组网（且物理链路和逻辑通道资源预设），实现网络与算力资源池之间的无缝对接。一方面，可根据接入算力资源的类型、规模及用户接入位置，按需升级或新增模块化网络组件，无需对整体架构重构，保障网络敏捷扩展；另一方面，能灵活适配不同层级自有及第三方 AIDC 的覆盖服务范围，实现异构算力的快速并网与池化管理，同时支撑网络长期平滑演进，避免因算力业务迭代导致的大规模网络改造，降低运维成本与架构调整风险。

b) 端网协同：支持 AIDC 内端侧与城域网络之间的深度协同，通过构建标准化、结构化的端网协同接口与信息交互机制，包括统一的数据模型、开放的南向/北向接口及低开销的实时遥测通道，实现业务状态（业务类型、优先级等信息）与网络状态（网

络拓扑、带宽利用率等信息）的双向感知、动态交互与联合决策，增强城域网络的拥塞避免、流量调度、QoS 协商等关键能力。

c) 广域无损：算力 POP 作为算网对接的关键节点，应具备与算力 POD 一致的广域无损能力，保障存算分离拉远训练、云边协同训推等算力业务的端到端 RDMA 流量传输。具体而言，算力 POP 需部署支持 400GE 高速端口和 GB 级大端口缓存的 DC-Leaf 设备，并具备精细化拥塞控制能力，有效避免网络丢包，确保 RDMA 流量的跨域高性能传输。

d) 确定性服务：算力 POP 需具备确定性服务能力，从而与其他网络组件协同，实现端到端的确定性服务。一方面，通过 FlexE、IP 确定性切片、VPN 等技术，面向多租户提供软硬结合的层次化网络切片服务，保障业务的 SLA 要求；另一方面，构建资源、路径、业务多层次确定性能力，在复杂业务负载和大规模流量并发场景下，为算力业务提供低时延、低抖动和高可靠数据传输服务。

7.3 算力互联区关键技术要求

算力互联区作为算力城域网与异构算力资源池及骨干网、核心网等外部网络互联的枢纽，可实现算力 POD、算力 POP 组件的灵活扩展和组件间流量的高效疏导。算力互联区可基于 400G/800G 大带宽链路、大象流负载分担等技术实现网络高吞吐，并可通过精细化拥塞控制、IP 确定性切片等机制实现各类业务端到端的无损、确定性保障。

算力互联区需满足如下关键能力要求：

a) 大带宽链路：为满足海量样本数据高速入算，以及算力资源整合带来的流量压力，算力互联区需支持 400G/800G 高速链路规模部署。通过大带宽链路承载聚合流量，有效提升带宽利用率和动态调度能力，构建高运力网络基础设施，持续降低单比特传输成本，并为未来向更高速率技术演进奠定基础。

g) 超高吞吐：由于智算业务中存在大量“流数少、流速高”的大象流，算力互联区需基于报文深度识别及 SRv6 流级调度技术，实现“大象流”的自动识别、子流拆分与精细化调度，最大化网络带宽利用率，保障网络吞吐率可达 90% 以上。

b) 高收敛比组网：算力互联区作为各组件互联互通的枢纽，需具备高收敛比组网能力，有效应对大规模数据并发同步与瞬时流量突发对网络造成的冲击。通过“突发缓存 + 队列调度”双重机制，一方面利用高速缓存模块吸收瞬时流量峰值，另一方面通过优

先级调度确保 GPU 控制信令、关键业务数据的优先传输，防止计算资源因等待数据而闲置。通过按需引入 8:1、16:1、32:1 等高收敛比，在保障算力传输效率的同时，实现建网成本与算效的最优平衡。

c) 广域无损：针对存算分离拉远训练、跨集群协同训练等算力业务场景中样本面、参数面数据长距离传递时的低丢包需求，算力互联区须依托 400GE/800GE 高速端口、端口大缓存及精细化拥塞控制机制，支持 K 级 RDMA 无损队列，有效抵御网络丢包，确保 RDMA 在跨域传输中仍保持高性能。

d) 确定性服务：算力互联区也需具备确定性服务能力，从而与算力 POD、算力 POP 协同，为入算、算间等各类算力业务提供确定性的低时延和稳定带宽：通过 FlexE、IP 确定性切片等技术，并结合 SRv6 可编程路径能力实现低时延算路与拥塞规避，确保各类算力业务的端到端差异化 SLA 保障；构建资源、路径、业务多层次确定性能力，确保业务在高负载情况下的低时延、低抖动和高可靠。

7.4 运营管理系统关键技术要求

运营管理系统是算力城域网的智能管控枢纽，该系统通过统一管控平台，对全域算力资源、网络资源及安全策略进行集中纳管与智能调度，支撑算、网的资源协同与业务联动。同时，系统可提供从资源分配、路径优化到 SLA 保障的自动化服务，实现故障自愈、安全主动防御与能效优化，为多租户提供可承诺、可度量、可运营的算网一体化服务，全面提升算力资源利用效率与业务体验保障能力。

运营管理系统需满足如下关键能力要求：

a) 弹性带宽：具备弹性带宽分配能力，支持根据用户业务周期、任务特征、QoS 需求，动态调整网络带宽资源，支持从 M 级到百 G 级带宽的按需灵活分配，避免带宽固定分配导致的资源闲置或拥塞，实现突发流量的敏捷响应，显著提升资源利用效率。

b) 任务式服务：支持基于任务式服务的带宽按需申领和动态分配机制，根据用户带宽需求（如最小保障带宽、峰值带宽、优先级等）自动配置端到端传输路径的带宽资源。此外，在流量突发或业务高峰期，可通过流量监控与预测机制实时感知负载变化，动态扩容或收缩带宽，保障高优先级业务的传输稳定性，显著提升网络与算力资源的整体利用率。

c) 网络级负载分担：当前的负载均衡技术基于 HASH 随机，只能做到流比较多时近似均衡散列，并不能保证所有链路都完美均衡。因此，运营管理系统需主动获取和解析 AI 流量通信关系，并基于全局视角对整网流量进行统一规划与调度，实现全网路径负载均衡效果最优，保障业务高吞吐，提升整网资源利用效率。

d) 网络自治：具备网络全域实时感知能力，实时监测拓扑、负载、设备健康状况及流量特征（如大象流、拥塞状况等），为调度提供依据。同时，可支持 AI 驱动的网络故障自愈，基于深度学习和知识图谱构建故障自诊断模型，实现根因推理与隐患识别，达成网络快速恢复。通过构建“感知 - 分析 - 决策 - 执行”全流程自治体系，推动网络自治层级从 L3 向 L5 演进。

e) 算网协同调度：构建基于地理位置、资源类型及实时负载状态的算力度量体系，实现对全域异构算力资源的统一量化与感知，为算力资源的灵活调度提供基础支撑。在此基础上，需建立智能算力调度体系，依据业务 SLA 要求动态匹配最优算力节点与业务路径，并按需分配切片、带宽等计算与通信资源，实时监测资源利用率和业务质量，形成端到端的任务式调度闭环，提升算网资源整体利用效率与业务响应敏捷性。

f) 精细化管控：以数字孪生技术与随流检测机制为基础，构建仿真预演、异常处置的精细化管控体系。依托数字孪生构建“镜像网络”，同步配置与流量特征，实现针对配置变更、路由调整等场景的精确仿真。支持端到端的网络性能实时检测，实现对关键业务流的时延、抖动、丢包等指标的逐跳精准测量，实现快速的异常波动捕捉与根因定位，有效提升网络故障预警及快速排障能力。

g) 主动安全防御：基于多维数据采集与 AI 智能分析实现潜在威胁识别、自动化策略调整：持续采集并分析流量行为、访问频率、协议合规性及用户权限变更等多维数据，再结合 AI 模型识别潜在威胁（如横向渗透、数据窃取、非法扫描），并进行自动化策略调整，如对高危源地址实施动态访问控制、对敏感业务流启用加密隧道或触发路径切换，构建主动防御能力。