



未来网络技术发展系列白皮书（2025）

算力城域网白皮书 （2025版）

第九届未来网络发展大会组委会

2025年8月

版权声明

本白皮书版权属于中国电信股份有限公司研究院及其合作单位所有并受法律保护，任何个人或是组织在转载、摘编或以其他方式引用本白皮书中的文字、数据、图片或者观点时，应注明“来源：中国电信股份有限公司研究院等”。否则将违反中国有关知识产权的相关法律和法规，对此中国电信股份有限公司研究院有权追究侵权者的相关法律责任。

编写说明

主要编写单位：

中国电信股份有限公司研究院

华为技术有限公司

全球固定网络创新联盟

中关村超互联新基建产业创新联盟

主要编写人员：

中国电信股份有限公司研究院：朱永庆、胡泽华、龚霞、袁世章、阮科、陈迅、杨冰、赖道宁、胡家元、张帆、皮谭昕。

华为技术有限公司：任广涛、曾毅、李智、赵浩宾、陈卓、于凤青、张潇潇。

中关村超互联新基建产业创新联盟：袁博。

前 言

2025 年初 DeepSeek 的爆火掀起了生成式人工智能的浪潮，带动大模型训练成本和推理成本的快速下降，驱动算力需求爆炸式增长。城域网络作为用户与算力资源间的关键桥梁，各类新兴算力业务对城域网络的网络架构、网络能力及服务模式等方面提出了新的要求。中国电信在 2024 年发布了《算力城域网白皮书》，首次提出算力城域网概念，获得业界的广泛关注，引领了城域网络发展新方向。伴随着产业生态与技术的发展，以及算力城域网研究与部署的深入，中国电信推出《算力城域网白皮书（2025）》，进一步明确城域网络在面向算力业务新场景、新需求下需具备的网络架构和关键技术能力。

本白皮书首先从算力产业发展、宏观政策以及服务模式等角度分析了算力发展态势，引出了算力城域网的概念；然后，针对算力业务需求展开分析，明确了算力城域网应具备的网络能力；其次，分析了算力城域网设计目标，阐述了算力城域网的总体架构、关键技术和设备能力；最后，给出了算力城域网面向具体业务场景的技术方案，以及演进路径建议。

本白皮书旨在通过对算力城域网的网络架构、关键技术、应用场景及发展策略的探讨，吸引更多行业内的专家和相关从业者共同参与算力城域网的创新发展与产品实现，推动网络向更加高效、智能、灵活、绿色、安全等方向演进。

目 录

前 言	I
目 录	II
一、引言	1
二、算力发展态势	2
2.1 算力产业发展和政策趋势	2
2.2 城市算力通过算力城域网就近服务本地用户	4
三、算力城域网需求	6
3.1 需求总述	6
3.2 网络能力需求	7
四、算力城域网总体架构	13
4.1 算力城域网架构设计目标	13
4.2 算力城域网总体架构	15
4.3 算力 POD	16
4.4 云网 POP	16
4.5 出口功能区	17
五、算力城域网关键技术	18
六、算力城域网设备能力要求	22
七、算力城域网典型应用	26
7.1 海量数据高效入算场景	26
7.2 存算分离拉远训练场景	27

7.3 跨集群协同训练场景	29
7.4 云边协同训推场景	30
7.5 推理下发场景	31
八、总结与展望	33
附录 A：术语与缩略语	34
参考文献	36

第九届未来网络发展大会白皮书

一、引言

数字经济已经成为全球经济发展的强大引擎，世界各国都在积极推进数字化经济新质生产力的建设。算力作为数字经济时代的核心生产力，正在加速数字经济和实体经济的深度融合。以 AIGC 为代表的人工智能大模型快速崛起，推动了算力需求的快速增长。

AI/HPC 等高阶算力对于提升国家、区域经济核心竞争力的重要作用已经成为业界共识。2025 年 3 月，《政府工作报告》提出将持续推进“人工智能+”行动，打造具有国际竞争力的数字产业集群。这意味国家将加强顶层设计，加快形成以人工智能为引擎的新质生产力。随着这一行动的深入推进，人工智能将在推动产业升级、促进新质生产力快速发展等方面发挥重要作用。

随着人工智能技术的快速发展，大模型训练对算力的需求呈指数级增长，促进了运营商、政府、行业和企业进行算力中心的建设。各类算力资源如何实现高效整合，服务于千行百业，进而实现算力的商业闭环是业界普遍关心的话题。本白皮书针对算力发展新态势和算力业务新需求，在 2024 年《算力城域网白皮书》基础上，对算力城域网的网络需求、网络架构、应用场景和关键技术等方面进行了更新和完善，旨在应对算力快速发展带来的各种挑战，实现城域范围内异构算力的资源整合和高效供给。

二、算力发展态势

2.1 算力产业发展和政策趋势

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。2025 年以来，以大语言模型 DeepSeek 系列为代表的人工智能，通过算法创新与工程优化解锁了更高的算力利用率，实现训练成本与推理成本的显著降低，加速人工智能的落地部署与普惠化发展，进一步催生了大量算力需求。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将达到 1037EFLOPS，预计到 2028 年将达到 2782EFLOPS，五年年复合增长率达到 46.2%。依托超大规模市场优势，我国算力水平和供给能力大幅提升，形成了体系完整、规模庞大的产业体系。

我国高度重视算力产业与技术的发展，中央及各级地方政府出台了一系列政策文件以引导其健康发展。例如，《数字中国建设整体布局规划》指出要系统优化算力基础设施布局，促进东西部算力高效互补和协同联动，并通过推动建设普惠算力、推动算力一体化服务等方式，降低中小企业用算成本。《数字中国建设 2025 年行动方案》提出开展“人工智能+”行动，深度挖掘应用场景，建设高质量数据集，目标到 2025 年数字经济核心产业增加值占 GDP 比重超 10%。《算力互联互通行动计划》指出集中力量开展高性能传输协议等网络传输技

术研究，推动数据通信产业高质量发展，加快高性能路由器、高速无损网络技术研究，支撑数据高效入算、算力无损互联。攻克算力标识关键技术，研制新型算力标识网关，提高多样化算力感知能力。

大模型是指具备大规模参数和复杂计算结构的机器学习模型，能够处理海量数据、完成各种复杂的任务，包含语言大模型、视觉大模型等多种类型。大模型的快速迭代极大地增加了智算算力需求。根据应用领域的不同，可分为通用大模型、行业大模型和垂直大模型等。通用大模型不受特定领域的限制，具备跨领域的泛化能力，需要更大的数据量和计算资源，训练成本普遍较高。行业大模型和垂直大模型专注于特定行业或应用场景，表现出更高的专业精度和深度，算力要求和训练成本也相对较低。大模型部署可分为训练和推理两种场景：训练场景是指使用大规模数据集对庞大、复杂的大模型参数进行训练的过程；推理场景是指在特定条件，利用大模型进行逻辑推理、知识推断提供高效准确结论的过程。

对于大部分企业和研究机构来说，自行购买和维护高性能计算设备训练大模型需要投入大量的人力和物力。在这个背景下，算力租赁服务应运而生。算力租赁通过即插即用的弹性解决方案，降低了算力获取难度和使用成本。企业的信息化和智能化转型可以通过算力租赁更快捷地获取所需的算力资源，从而加速技术研发和创新。当前，算力租赁服务已成为主流解决方案，不仅降低了算力使用的门槛，推动了算力技术的普惠化发展，更改变了算力资源的配置方式，为算力服务的广泛应用和持续创新提供了强有力的基础设施支撑。

2.2 城市算力通过算力城域网就近服务本地用户

面对算力产业快速发展趋势，中国电信推出了面向智算的新一代数据中心（AIDC，Artificial Intelligence Data Center）基础设施技术方案，在高算力规模、高功率密度、高弹性需求背景下，同时兼容通算、超算、智算，甚至未来的量子算力。在 AIDC 基础上，中国电信积极布局算力基础设施建设，面向政府、企业、科研机构等提供就近的高效算力服务。除了围绕热点区域建设“2+3+7”¹的公共智算云池，还重点在各省打造（N 个）城市智算基础设施服务当地用户，满足数字政务、城市大脑等城市数字化转型场景；通过（M 个）行业智算集群服务教育、电力、金融、互联网等重点行业；并基于地市边缘节点（X）按需将推理算力向边缘覆盖。

中国电信云化 IP 城域网(简称为新型城域网)具备原生算力业务高效承载的能力，基于云网 POP 灵活架构以及城域 Spine-Leaf 的 Full-Mesh 组网优势，实现了云边/边边高效协同和算网快速对接。面向算力业务的长期演进，中国电信通过引入算力灵活调度、算力无损传输、精准流级调度、网络智能运维等能力，打造以算力为中心、算网一体的城域网新业态——算力城域网²。当前，中国电信在上海、浙江、广东等地围绕海量数据弹性高效入算、存算分离百公里拉远训练、百公里分布式推理等新型智算业务开展算力城域网现网验证，实

¹ “2”是指中国电信在内蒙古、贵州设置的两大公共训练智算中心；“3”是指中国电信在北京、上海、广州设置的三大热点区域智算中心；“7”是指中国电信在安徽、宁夏、湖北、江苏、辽宁、重庆、浙江设置的七大区域训练智算中心。

² 算力城域网：Computing service Oriented Metropolitan Area Network (COMAN)。

验结果表明算效、TTFT、TPOT 等指标劣化小于 3%，充分证明算力城域网方案的技术可行性。

算力城域网具备算力高效整合、算力无损输送、算力服务即取即用等关键能力，通过构建 AIDC 与用户之间的安全高速通道，支撑城市算力和行业算力就近服务本地算力用户。算力城域网通过高弹性、高吞吐、高可靠的一跳入多算等网络新型服务能力，为政府、企业、科研机构各类客户提供高效便捷的算力服务，加速数字化转型进程，支撑数字经济的高速发展。

第九届未来网络发展大会白皮书

三、算力城域网需求

3.1 需求总述

典型算力业务流程包含数据入算、模型训练、模型推理几个关键阶段。数据入算阶段要求网络能够满足 TB/PB 级海量数据的高效传送。考虑到数据敏感用户的数据“不落盘”要求，还需要网络具备高稳定和数据无损传输能力，实现用户私域存储与 AIDC 之间的高效拉远训练。模型训练阶段当前面临单 AIDC 算力资源受限、零散算力资源未利用等问题，亟需通过分布式协同训练实现算力资源高效整合，要求网络提供无损、高吞吐的高性能算间互联。模型推理阶段包含推理结果生成和推理结果下发两个关键步骤：推理结果生成需要大量算力资源以保证海量用户并发推理体验，网络需具备无损传输、高可靠能力，满足分布式推理需求；推理结果下发需要保障用户的泛在接入与实时交互，要求网络具备广覆盖及确定性服务能力。

算力城域网作为连接用户与算力资源的关键桥梁，为算力租赁服务提供了关键的网络支撑，确保租用算力资源的企业可以获得接近本地部署的算力使用体验，需要高效满足海量数据入算、存算分离拉远训练、跨集群协同训练、云边协同训推、推理下发等算力业务需求。

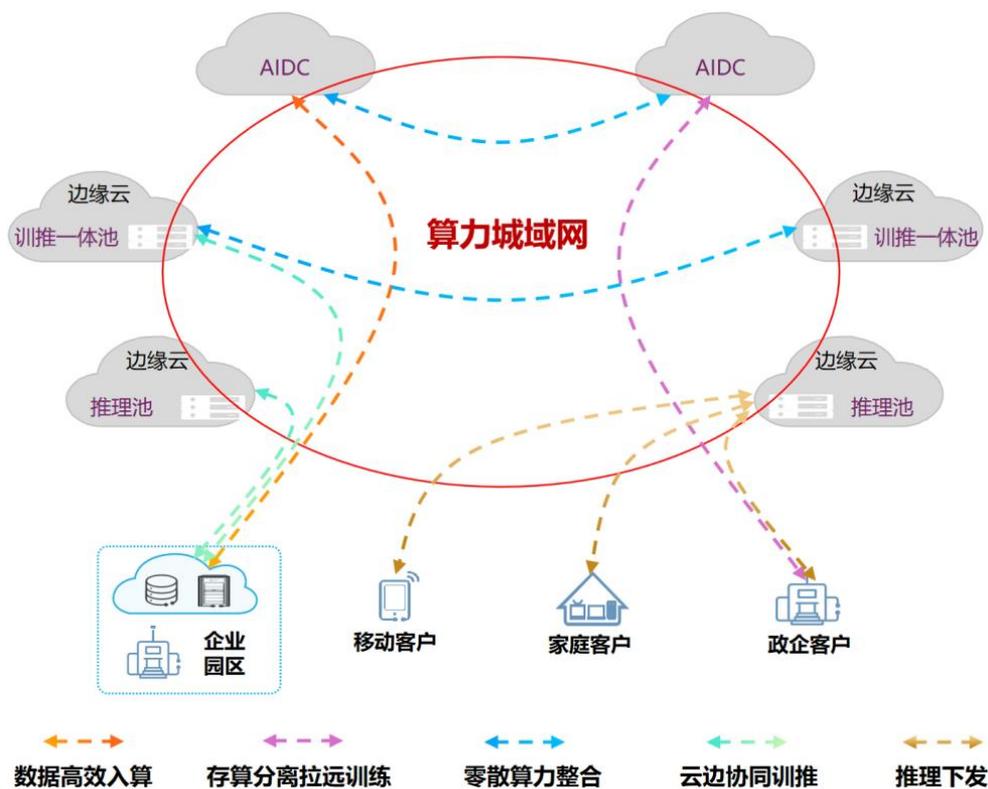


图 3-1 算力城域网业务需求总览

3.2 网络能力需求

3.2.1.海量数据高效入算需求

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到数百 TB 的量级。根据 IDC 发布的《Worldwide IDC Global DataSphere Forecast Update,2025 - 2029:Regional Analysis》报告，IDC 预计，2025 年全球将产生 213.56ZB 数据，到 2029 年将增长一倍以上达到 527.47ZB；其中，中国市场 2025 年将产生 51.78ZB 数据，到 2029 年增长至 136.12ZB，CAGR 达到 26.9%。众多企业当前仍依靠邮寄硬盘的方式进行大规模数据的搬运，诸如科研、交通、影视、医疗等领域，每年都有 PB 级数据通过硬盘

搬运/邮寄方式传送到算力中心进行模型训练。这种“人工拷贝+硬盘快递”的模式不仅效率低，还面临着硬盘损坏与数据丢失的风险。当前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。

为了实现海量数据的高效流转，网络需要具备网络级负载均衡能力，全面提升整网的有效吞吐量，打造高运力网络，为数据高效传送提供坚实基础。同时，网络需要具备高度的弹性与敏捷性，基于任务式服务为企业按需提供弹性带宽，满足短时间内大批量数据传输的需求。

总之，海量样本的快速入算服务对算力城域网的需求是：实现 TB/PB 级数据的弹性带宽服务，时间可承诺（分钟达、小时达、天级达）；业务分钟级开通，任务式服务。

3.2.2.存算分离拉远训练需求

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，医疗领域涉及的电子病历、流行病监测数据以及基因序列分析等敏感信息，这些数据均具有较高的安全标准。在算力租赁场景中，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。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，算力资源节点与样本数据存储节点需要跨广域部署，并且在模型训练时需要保持频繁的实时交互，以分批拉取所需的样本数据。

在此场景下，由于样本数据传输采用对时延、丢包高度敏感的 RDMA 协议，网络除了要具备高弹性、高吞吐能力外，还需要具备 RDMA 无损传输能力，以确保模型训练的高效性和稳定性。此外，网络还需要部署强健的数据加密机制，保障样本数据传输的安全性。

综上，存算分离拉远训练服务对算力城域网的需求是：实现用户私域存储到 AIDC 之间 100km-500km 的高效拉远训练，数据广域无损传输保障算效下降小于 5%；支持拉远训练过程中的数据安全隔离保障；网络链路和资源能够达到 90%以上高吞吐能力，充分提升网络带宽资源利用率。

3.2.3.跨集群协同训练需求

大模型 Scaling Law 持续生效，十年间大模型的算力需求增长约 100 万倍，后续大模型的算力需求预计仍将保持每年 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 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，整合零散的社会算力有助于打破地域、机房资源、服务商等限制，构建统一、高效的算力服务平台。

跨集群协同训练对网络带宽、时延要求高，网络需具备 400G/800G 大带宽链路以及 RDMA 无损传输能力，避免算力传输过程中出现拥塞、丢包。同时，考虑到训练参数同步时流量大并发、高突发特性（以万卡级 AIDC 为例，单网卡 200Gbps 的传输速率将导致参数面突发流量峰值高达 2000Tbps），需按照最优收敛比建网，平衡算效与建网成本。此外，在大模型训练过程中，一旦发生网络故障引发训练任务卡死等问题会严重影响训练效率，网络须具备高精仿真、网络自愈等智能运维能力。

综上，跨集群协同训练服务对算力城域网的需求是：采用 400G/800G 高速链路，支撑 100km-500km 跨集群协同训练。基于 RDMA 无损数据传输保障跨集群训练的算效下降小于 5%。采用 4:1、8:1、16:1、32:1 等高收敛比组网；网络高稳定运行，故障影响不扩散。

3.2.4.云边协同训推需求

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少量训推一体机实现大模型的快速应用。但是，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。因此，通过企业本地算力与云端租赁算力之间的高效协同，以满足企业算力资源灵活扩展需求，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，并结合输入、输出层的本地化部署，保证样本数据不出园区，满足了金融、医疗等数据敏感客户的需求。

数据安全要求。此外，该方案与 Prefill-Decode 分离式推理架构天然适配，通过将可弹性扩展的 Prefill 和 Decode 分层部署在本地和云端，显著提升推理效率、资源利用率和用户体验。

在此场景下，本地和云端算力池之间基于 RDMA 协议传递参数面同步数据、KV Cache 等信息，要求网络在保证高吞吐、低时延的前提下，具备 RDMA 无损传输能力，避免出现拥塞、丢包。同时，网络还需通过租户级切片保障业务间的有效隔离，在保障业务 SLA 要求的同时，避免业务之间故障相互影响。

综上，云边协同训推场景对算力城域网的需求是：具备网络级负载均衡和 RDMA 无损传输能力，保证在用户吞吐不下降的同时，TTFT 和 TPOT 劣化低于 5%；具备端到端租户级业务隔离能力，有效保障多租户 SLA。

3.2.5.推理下发需求

推理应用的规模落地是大模型技术实现商业变现的核心环节。华为预计 2025 年 70% 的新应用将集成 AI 模型，端侧设备（如智能手机、智能家居设备等）与推理池之间的并发量将达到百万级。端侧设备与推理算力集群间的高频、富媒体即时交互，对网络的响应速度与带宽提出了更高的要求，以确保用户体验推理业务的流畅。

为了满足推理下发对低时延、高带宽及确定性的需求，网络需要实现泛在覆盖与便捷接入，确保用户能够享受一致的服务体验。同时，网络还需具备确定性服务能力，能够精准识别并优化数据传输路径，

提高数据传输的确定性和可靠性，从而满足用户的实时交互需求。

综上，推理场景对算力城域网的需求是：泛在接入各种算力用户，满足百万级用户的并发带宽需求；具备确定性服务能力，对于时延敏感类业务确保 RTT 小于 10ms；对数据敏感的行业和政企用户，需要基于网络切片实现推理业务数据的安全隔离。

第九届未来网络发展大会白皮书

四、算力城域网总体架构

算力城域网是中国电信从云网协同到云网融合战略的具体实践，通过“算力”、“云”和“网”在基础设施层的融合，实现“算力”、“云”和“网”在逻辑架构、资源管理和服务调度方面的逐渐融合，支撑云网融合的算网产品和服务的持续创新。算力城域网依托新型城域网的架构优势，构建以算力为中心、算网一体的新服务、新平台、新形态。

4.1 算力城域网架构设计目标

（1）凝聚算力，共筑生态

- 联接智算、通算、超算、量子、安全等数字能力，实现算网多要素融合、多能力一体化服务。
- 联接 DC、自有云、三方云、社会算力、IT 生态，构筑丰富的业务和应用生态。
- 联接 ToC/ToB/ToH 场景下的海量用户资源，将算力服务和生态应用引入到千行百业、千家万户。

（2）算网一体、灵活部署

- 沿用新型城域网模块化组网架构，基于 Spine-Leaf 灵活扩展，泛在接入各类用户。
- 算力网关随 AIDC 灵活部署，构建网络和算力资源的标准化对

接模型，实现网随算动。

- 基于 SRv6/EVPN/网络切片等 IPv6+技术底座，实现对 RDMA 等数据传输协议的统一承载。

（3）算网赋能，使能商业

- 引入弹性带宽、超高通量、广域无损等新技术，支撑存算分离拉远训练、跨集群协同训练等创新业务和服务。
- 基于大象流自动识别与智能调度，实现网络级智能负载均衡，达到全网资源利用率最优，提升投资收益比。
- 通过算力业务应用感知和流级精细化调度，支撑差异化算网产品和服务的商业创新。

（4）智能运维、安全可靠

- 实现高精仿真，消除因配置差错导致的网络事故。
- 打造精细化业务监测能力，实现全网资源与业务实时可视。
- 基于网络实时监测和故障快速感知，实现故障自动定界定位。
- 按用户或业务类型设置网络切片，实现用户数据管道安全隔离。

4.2 算力城域网总体架构

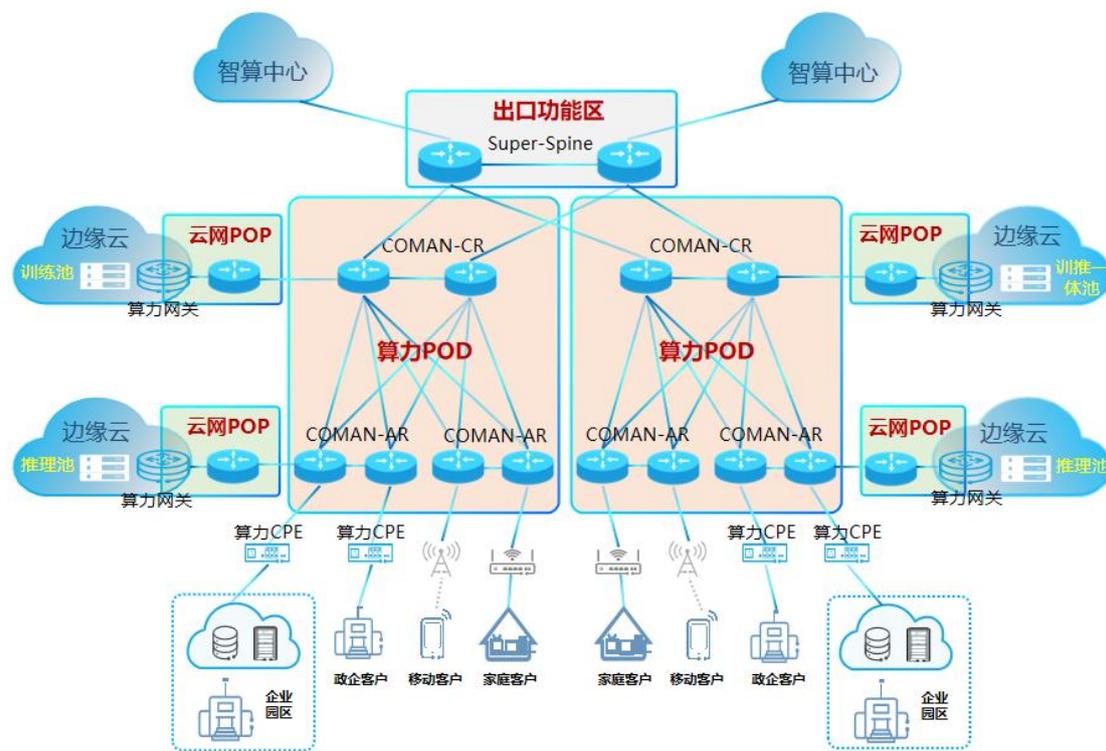


图 4-1 算力城域网（COMAN）总体架构

算力城域网基于算力 POD、云网 POP、出口功能区模块化组件搭建。算力 POD 实现 ToC/ToB/ToH 场景的算力用户统一接入和广泛覆盖，以及深/浅边缘算力池的快速接入；云网 POP 通过算力网关联接城域网络与算力资源池，实现二者在参数面、样本面、业务面网络间的标准化快速对接；出口功能区作为 AIDC 与算力 POD 之间的枢纽，实现用户数据高效入算，以及多 AIDC 之间的算力协同。

算力城域网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，通过与云网 POP 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，实现跨 POD/跨域的用户至算力资源池的端到端无

损连接。三大模块均采用 SRv6/EVPN 等标准化技术底座，在确保端到端业务逻辑一致性的同时，为算力业务提供高质量的网络承载能力。

4.3 算力 POD

算力 POD 可根据用户位置、行政区域、AIDC 服务范围等因素灵活设置，实现算力用户终端、企业分支站点的融合接入，涵盖光纤、PON、5G 等多种接入介质。算力 POD 基于算力 SPINE 设备（COMAN-CR）、算力 Leaf 设备（COMAN-AR）设备构建的 Spine-Leaf 架构，实现固、移、云、算业务的统一接入与融合承载，同时保证了区域内各类业务流量的无阻塞快速转发。算力 POD 可为用户提供无损的算力通道，实现算力服务的广覆盖和高效输送；并结合云网 POP 快速接入深/浅边缘推理池，实现边缘算力的池化利用和灵活调度，为用户提供低时延、高体验的算力服务。

4.4 云网 POP

云网 POP 内部署 COMAN-AR 与算力网关对接，通过城域网络与算力资源池的标准化、快速对接，实现“小时级”算网架构互联、“分钟级”算网业务开通以及 ms 级算网业务访问时延，支撑算网一体化服务。云网 POP 作为算力资源池的网络锚点，可接入自有及第三方异构算力资源池，实现算力资源池化与统一管理；同时通过与省级或区域 COMAN-CR 设备互联，实现跨域算力资源池之间的低时延、高可靠连接，打通跨域算力协同网络。

4.5 出口功能区

出口功能区作为城域网与骨干网、互联网及行业专网的枢纽，简化了城域网与外部网络、异构算力资源池之间的连接，实现组件灵活扩展以及组件间流量高效疏导。其中，Super-Spine 做为流量汇聚节点联接多个算力 POD，通过弹性带宽、网络级智能负载均衡等技术承载样本数据高速入算、存算分离拉远训练等业务；Super-Spine 同时实现多 AIDC 互联，基于 400G/800G 高速链路、广域无损等技术实现城域网内异构算力的整合和协同调度，支撑算力分布式训练或推理场景。

构建用户到 AIDC 以及 AIDC 之间的高吞吐、无损互联网络是算力城域网解决方案的设计重点。算力城域网通过引入弹性带宽、高吞吐调度、高收敛比组网、广域无损、精准流控、智能运维等网络技术和能力，实现用户到 AIDC 入算业务以及 AIDC 之间联算业务的高效承载。同时，算力城域网提供租户级网络切片等软/硬隔离技术，为用户提供安全隔离、确定性保障、高可靠的入算新服务。

五、算力城域网关键技术

算力城域网将算力服务由数据中心机房向广域网延伸，必须具备相应的关键技术和能力。

（1）弹性带宽能力

企业普遍面临周期性数据传输带来的带宽配置难题：长期采用高带宽专线会导致闲置期资源浪费，而低带宽专线则因传输时延过长造成算力资源空转。算力城域网需具备弹性敏捷的带宽按需拆建能力，可根据用户算力需求的变化动态分配带宽资源，为用户提供 100Mbps 至 100Gbps 的弹性带宽服务。根据企业算力业务需求和运营商网络资源使用情况，算力城域网可动态调度算力业务流的网络路径，在网络数据流量波动时仍能提供持续稳定的数据传输服务。

（2）超高吞吐能力

在智算、超算等场景中常常存在大量的大象流，这类流量具有单流瞬时高速率（10M-100G）的特征。传统基于五元组哈希的负载均衡技术难以有效应对大象流，容易造成负载不均衡和网络拥塞。算力城域网引入大象流自动识别与拆分、网络级负载均衡等关键技术，基于 IB 传输层信息（如 QP 对等）将大象流拆分为多个子流，并结合智能管控实现全域流量均衡，确保网络链路达到 90% 以上的超高吞吐率。在确保 AIDC 算效高效发挥的同时，实现算力城域网整网带宽和容量资源的高效利用。

（3）广域无损技术

对于存算分离拉远训练、跨集群协同训练以及云边协同训推等场景，AIDC 的样本面和参数面数据跨广域传输距离可达百公里甚至千公里。这两种场景都需要采用 RDMA 作为承载协议，提升数据传输的效率。RDMA 对网络性能非常敏感，0.1%的丢包将导致数据吞吐率下降 50%。算力城域网通过引入 400GE/800GE 端口和端口大缓存，有效降低网络拥塞概率，并结合流级拥塞控制机制，避免算力数据传输过程中发生丢包。确保跨广域训练的算效相较于单 DC 仍能达到 95%以上，并且云边协同推理相比集中推理 TTFT 和 TPOT 性能劣化小于 5%。

（4）收敛组网技术

在跨集群协同训练场景下，城域网络需要承载多个 AIDC 间的大规模参数面数据并发同步。算力城域网需具备收敛组网能力，有效缓解多 AIDC 协同训练时数据并发和突发对网络的影响：通过分层梯度聚合算法重构集合通信流程，有效减少跨数据中心通信的算卡数量，实现网络带宽的初步收敛；再采用“突发缓存+队列调度”的双重机制，利用高速缓存吸收流量冲击，并通过优先级调度确保 GPU 控制信令的及时传输，避免计算资源闲置等待。按需引入 4:1、8:1、16:1、32:1 等网络收敛比，实现建网成本与算效的最优平衡。

（5）模块化组网能力

算力城域网以 AIDC 为中心组网，基于模块化组网架构与 AIDC 标准化对接，实现超算、智算等异构算力的并网与池化调度，灵活适

配不同层级的自有及三方 AIDC 的覆盖服务范围。同时，算力城域网可根据算力用户接入位置和业务类型，按需升级或新建模块化网络组件，在敏捷、灵活、泛在接入各类算力用户的同时，保证网络平滑演进。

（6）一线接入、融合承载能力

算力城域网需构建基于统一协议栈同时承载固、移、云、算多元化业务的能力，在有效降低网络复杂度的同时，显著提升业务部署和运维效率；并基于 IP 的灵活连接能力优势，支持企业和行业用户一点接入，避免开通多种业务时部署多种 CPE，实现一个盒子一根线同时入网、入云、入算的便捷服务。

（7）超高可靠能力

算力城域网应保证长期稳定，防止因为链路拥塞、光模块故障、光纤质量下降等故障导致训练的中断。算力城域网需具备租户级故障隔离能力，实现租户级精准反压，避免故障扩散影响多租户算效。同时，算力城域网需具备随流检测、高精仿真、网络自愈等智能运维能力，通过业务流级可视、秒级流量趋势展示、秒级故障感知和逐包故障定界等技术，实现业务流故障的快速定位与恢复，驱动网络自治等级从 L3 有条件自治向 L4 高级自治演进。

（8）超高安全能力

算力城域网需要构建多层次的安全防御体系，应对多租户环境下的数据泄露和横向渗透风险。其核心在于实现租户数据的全流程隔离与端到端加密传输。算力城域网可基于 VPN 与网络切片技术，构建

“接入设备-网络切片-VPN”三级隔离机制，通过物理层、协议层和业务层的全维度解耦，有效阻断安全威胁。同时，算力城域网可通过结合量子加密、区块链等前沿技术，进一步增强网络的安全防护能力，向“主动免疫、动态感知、全链可信”的零信任架构演进。

第九届未来网络发展大会白皮书

六、算力城域网设备能力要求

算力城域网设备包含算力网关、算力 SPINE 设备（COMAN-CR）、算力 LEAF 设备（COMAN-AR）和算力接入 CPE 设备。算力网关是连接算力城域网与算力资源池的枢纽；COMAN-CR、COMAN-AR 是算力城域网的综合业务承载核心；算力 CPE 设备是算力用户的网络起点。这些设备作为算力城域网的核心组成部分，在硬件与软件能力上需满足以下要求，以实现超高吞吐、广域无损、超高可靠以及智能运维等关键能力。

（1）高速接口

当前，网关设备接口以 10GE 为主，辅以少量 100GE 接口；城域 SPINE 设备以 100GE、50GE 为主，辅以少量 10GE 接口；城域 Leaf 设备接口以 50GE、25GE、10GE 接口为主，辅以少量 100GE 接口；CPE 设备以 10GE、GE 为主，辅以少量 100GE 接口。

为了适应算力业务高吞吐需求，设备需升级端口能力：算力网关设备需支持单端口 100GE，且具备 400GE 演进能力；COMAN-CR、COMAN-AR 设备需支持单端口 100GE、400GE，且具备 800GE 和 1.6TE 演进能力；算力 CPE 设备需支持单端口 10GE、100GE，且具备 400GE 演进能力。

（2）动态大缓存

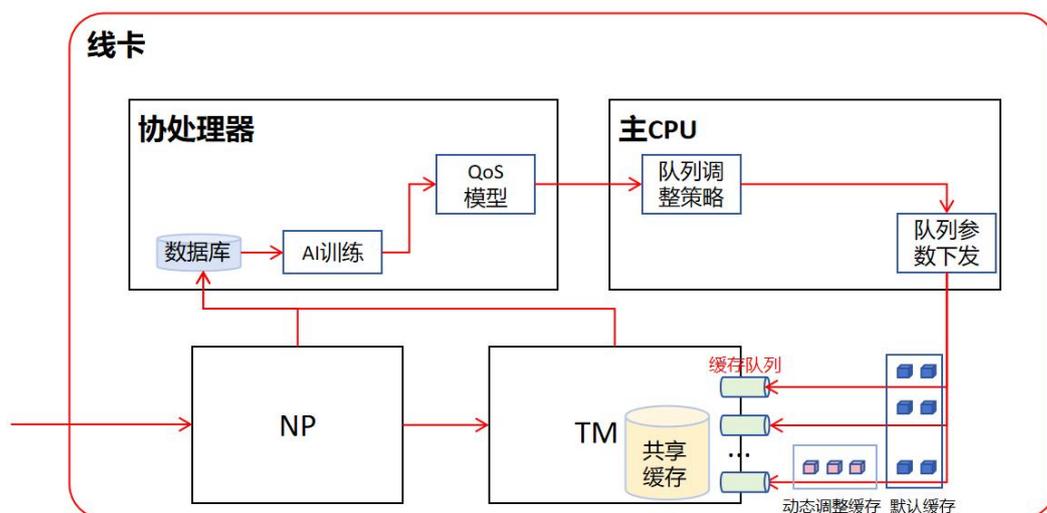


图 6-1 动态缓存机制示意图

针对广域 RDMA 高吞吐、高突发的特点，算力城域网设备需具备 GB 级端口大缓存，有效降低网络拥塞概率：一方面平滑因多流汇聚导致的瞬时速率波动；一方面为 PFC、ECN 等流控机制提供足够的反应时间窗口。此外，设备需具备缓存队列动态调度能力，实时感知业务流量变化以优化缓存分配策略，实现物理资源与业务需求的弹性匹配，支撑基于 SLA 的精细化业务保障。

（3）大象流拆分及负载分担

单条大象流带宽大、历时长，在与普通数据流混合转发时，极易出现部分链路拥塞、部分链路空闲的现象。这不仅导致网络全局负载不均，而且导致因拥塞而出现的 RDMA 报文丢失现象。为此，在基于五元组识别流的基础上，算力城域网设备需支持识别 RDMA 报文的 IB 传输层信息（如 QP 对、Partition Key 等），将单条大象流根据 IB 传输层信息划分多条子流。以此为基础，结合 Hash 算法优化实现基于子流的大象流负载分担，最大化网络吞吐与链路利用率。

（4）流量实时统计与上送

为实现流量的精细化管控，算力城域网设备需具备流状态表，通过对流/子流的速率、类型、优先级、QP 对等特征的实时统计，识别需重点保障的业务流和控制流。再基于 Telemetry 技术周期上报功能将流信息上送到控制器，由控制器为关键流定制最优路径，从而为业务流提供高吞吐、低时延、低拥塞的传输环境，并保障控制流的可靠、优先传输。此外，设备在上送流信息时可同时上送资源信息（可用带宽、缓存队列深度等），支撑控制器统一规划整网流量，实现全网路径间的无冲突均衡调度。

（5）流级拥塞控制

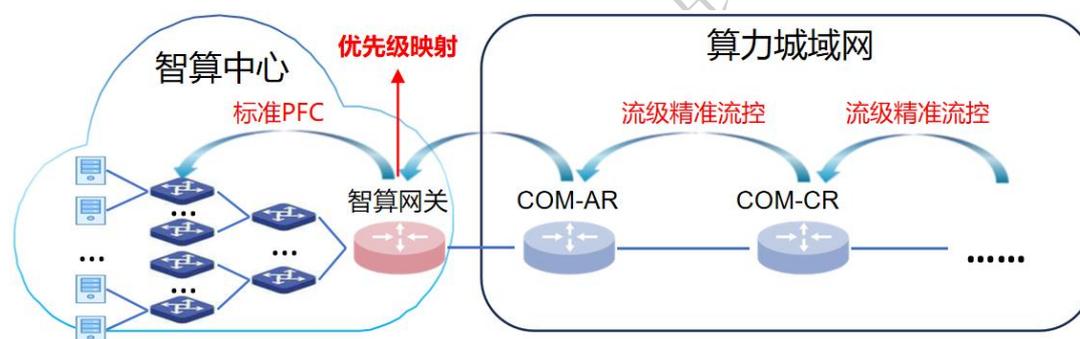


图 6-2 端到端精准流控表示意图

数据中心普遍采用 PFC 机制保障无损传输，但其端口队列级的粗粒度控制容易引发头阻和误伤问题。算力城域网融合承载固、移、云、算多种业务，因此更需要具备端到端的流级精准调控能力，避免因传统拥塞控制机制在广域场景下的滞后与扩散问题带来的吞吐下降。算力城域网设备需为每个 RDMA 业务流分配独立的缓存队列并实时监控缓存水位，当队列深度超过预设阈值时，基于端到端/逐跳

回溯方式向上游设备发送流级反压信号，从而实现细粒度的流量控制，在有效避免拥塞发生的同事，也避免了对其他业务的影响。此外，算力网关需支持流级精准流控技术与 PFC 之间的优先级映射，实现跨广域端到端的精准流控。

第九届未来网络发展大会白皮书

七、算力城域网典型应用

算力城域网是云化 IP 城域网面向算力新业态的演进，其组网方案和能力规划需要考虑算网一体化及算力业务创新的长期发展,根据业务需求，引入弹性高吞吐、高收敛组网、广域无损等能力，以应对海量数据高效入算、存算拉远训练、跨集群协同训练等新型算力业务场景。

7.1 海量数据高效入算场景

数据高效入算主要解决企业租赁智算资源时，TB/PB 级训练样本周期性上传的问题。算力城域网需要提供高效、敏捷、智能的大数据传输服务，并具备分钟达、小时达、天级达的多级服务能力。该场景组网方案如图 7-1 所示。

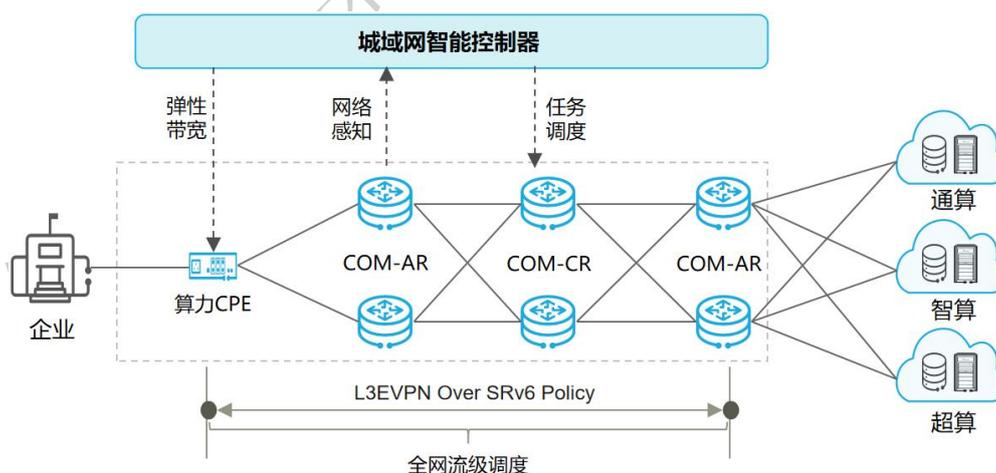


图 7-1 快速入算场景组网方案

智算 CPE 至入算网关（S-Leaf）之间采用 L3EVPN over SRv6

Policy 技术承载样本快速入算业务。结合任务式弹性带宽技术，实现用户入算带宽的弹性伸缩。城域网智能控制器可实时感知全网状态，实现基于任务流的全网调度，提高整网效率。对算力城域网的关键技术要求如下：

- **网络级负载均衡：**基于全局视角实现高效、动态且均匀的流量分配，保障端到端流量全局负载均衡，提升网络吞吐能力。
- **弹性带宽服务：**实现 100Mbps 至 100Gbps 的弹性带宽分钟级开通与秒级变更，根据业务需求灵活调整带宽容量。
- **一线入多算：**全网任意节点间 L3 层连接可达，确保企业侧数据一线接入通算、智算、超算等多数据中心。
- **高可靠性：**基于高精仿真、故障自愈等技术，实现网络 0 事故、业务 0 中断，确保数据高速传输的连续性和稳定性。

7.2 存算分离拉远训练场景

存算分离拉远训练主要满足企业数据敏感时，样本数据不在园区外“落盘”的拉远训练需求。存算拉远训练要求实现城市内、省内或区域范围内的数据高效传输和训练，保障训练效率和数据安全。算力城域网需具备高效、安全、智能等能力，保障训练效率和数据安全。该场景组网方案如图 7-2 所示。

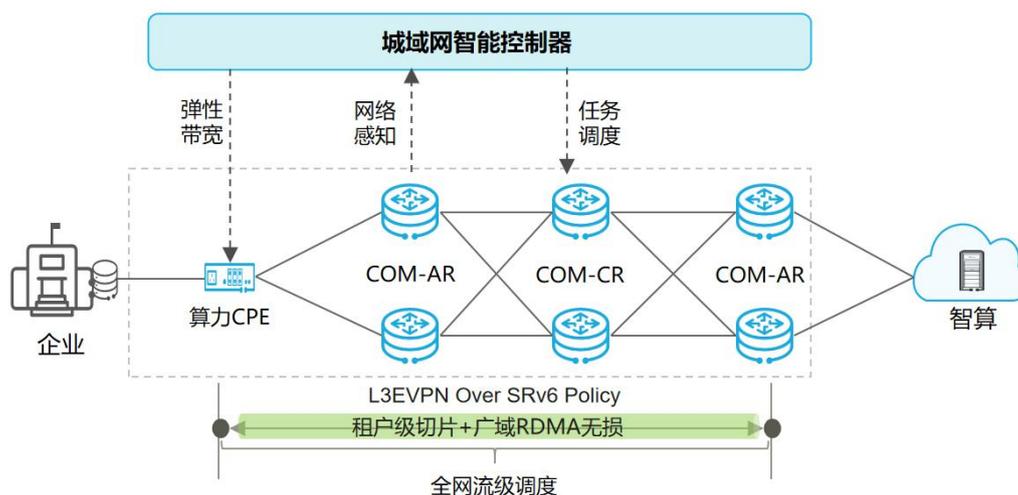


图 7-2 存算拉远场景组网方案

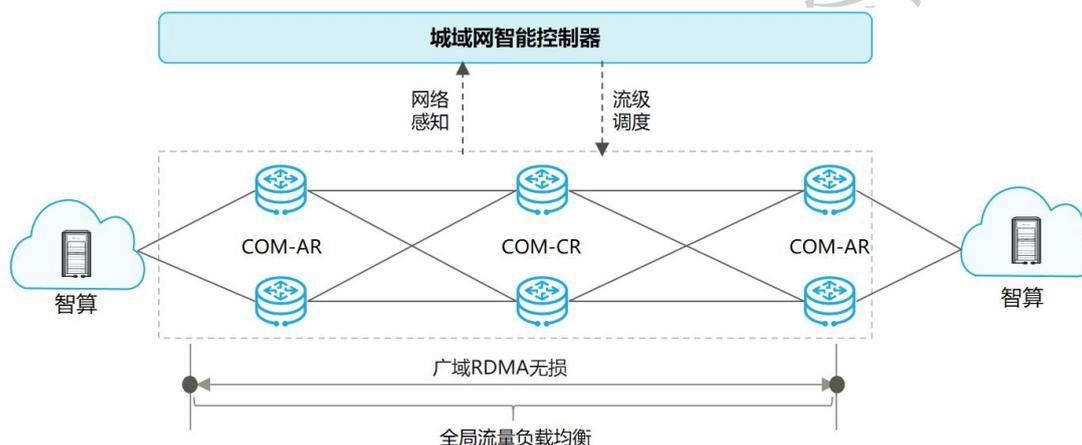
智算 CPE 至入算网关（S-Leaf）之间采用 L3EVPN over SRv6 Policy + 租户级切片承载存算分离拉远训练业务。通过 RDMA 广域无损技术保障数据传输的高效性和稳定性。同时，在网络边缘节点通过访问控制策略等安全防护措施实现业务安全隔离，满足数据敏感企业的需求。对算力城域网的关键技术要求如下：

- 网络级负载均衡：基于全局视角实现高效、动态且均匀的流量分配，保障端到端流量全局负载均衡，提升网络吞吐能力。
- 弹性带宽服务：实现 100Mbps 至 100Gbps 的弹性带宽，根据业务需求灵活调整带宽容量。
- 广域无损保障：通过智能流量调度和拥塞控制等技术，保障数据传输的高效率和稳定性，实现 RDMA 报文广域无损传输。
- 多租户故障隔离：租户级切片+租户级精准流控，实现租户间业务隔离，保障租户业务故障影响不扩散。
- 高可靠性：基于高精仿真、故障自愈等技术，保障存算拉远训练的算效。

- 业务智能：全网资源/业务可视，基于网络传输能力实现任务流级调度，保障训练效率。

7.3 跨集群协同训练场景

跨集群协同训练旨在高效整合多个 AIDC 的算力资源支撑大模型训练。跨集群协同训练要求在保证算效的前提下，通过合理的带宽收敛降低多 AIDC 组网的成本，实现训练效率和建网成本的最优平衡。该场景组网方案如图 7-3 所示。



多个联算网关（S-Leaf）之间采用 L3EVPN over SRv6 Policy 技术承载跨集群协同训练业务。基于大带宽链路和广域 RDMA 无损能力，保障多 AIDC 协同训练算效。通过端口缓存等技术实现高收敛比建网，降低模型训练的带宽需求。对算力城域网的关键技术要求如下：

- 大带宽链路：端到端 400G/800G 高速链路，满足协同训练的高带宽需求。
- 广域无损保障：通过智能流量调度和拥塞控制等技术，实现 RDMA 报文广域无损传输。

- 高收敛比组网：通过端口缓存等技术实现带宽收敛，降低训练过程中的并发与突发对网络带宽的需求，节省建网成本。
- 高可靠性：基于高精仿真、故障自愈等技术，保障多 AIDC 协同训练的算效。
- 智能运维：通过业务流级可视、逐包故障定界等技术，实现业务流级的故障定界定位。

7.4 云边协同训推场景

企业本地部署训推一体机，通过租赁运营商智算池的方式应对企业快速增长的模型微调、业务推理对算力需求。通过训推一体机与算力资源池之间的云边协同训练/推理，实现企业算力资源的弹性扩展，云边协同训练/推理基于模型分割实现，这种方式既满足了企业推理/训练敏感数据不出园区的要求，又满足企业算力的灵活扩展按需租赁云端算力的诉求。由于城域网需要支撑层间的参数面数据同步因此需要具备 RDMA 无损传输与超高吞吐能力。该场景组网方案如图 7-4 所示。

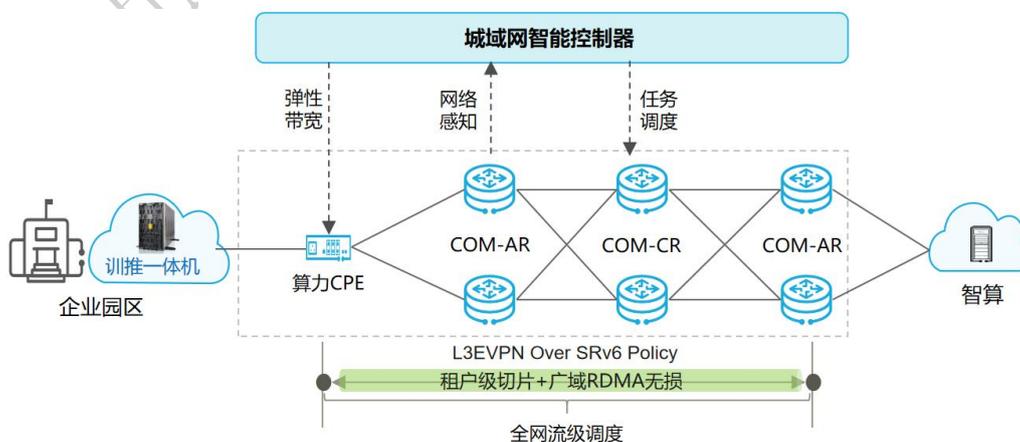


图 7-4 云边协同分布式训推组网方案

智算 CPE 至入算网关（S-Leaf）之间采用 L3EVPN over SRv6 Policy + 租户级切片承载云边协同分布式训推业务。通过 RDMA 广域无损技术保障数据传输的高效性和稳定性。对算力城域网的关键技术要求如下：

- 网络级负载均衡：基于全局视角实现高效、动态且均匀的流量分配，保障端到端流量全局负载均衡，提升网络吞吐能力。
- RDMA 广域无损：通过流级精准流控，避免在模型训练的过程中发生丢包，保障边侧算力与云侧算力协同过程中算效不下降。
- 租户级切片隔离：支持将云边协同分布式训推流量与普通业务流量隔离开，避免拥塞控制过程中的流量压制而对其他业务产生影响。
- 高可靠性：基于高精仿真、故障自愈等技术，保障云边协同分布式业务可用性。

7.5 推理下发场景

推理业务主要为企业基于业务诉求的差异化推理服务和算力调度。其要求网络提供差异化的业务承载和算力调度能力，满足百万用户并发的确定性时延需求。网络需具备广覆盖接入、切片隔离、数据加密等能力，保障推理业务的连续性和安全性。该场景组网方案如图 7-5 所示。

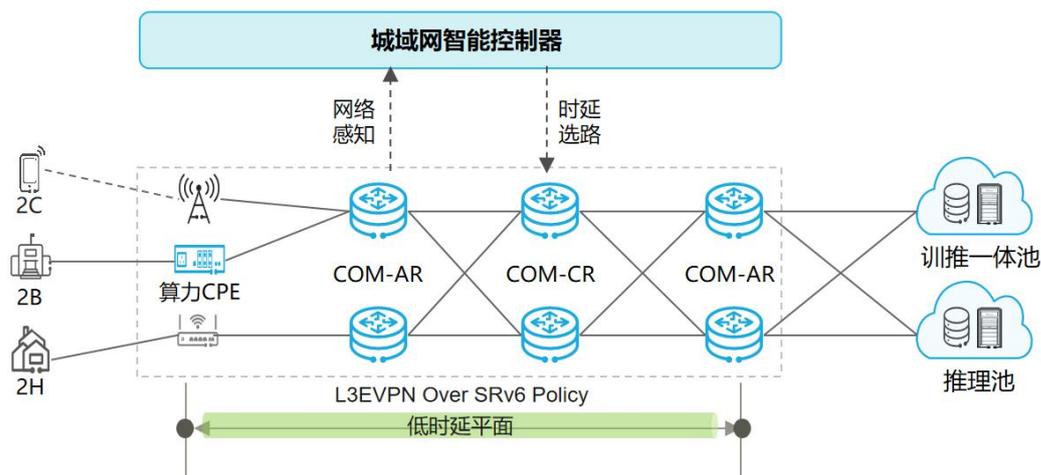


图 7-5 推理下发场景组网方案

智算 CPE 至推理网关（S-Leaf）之间部署 L3EVPN over SRv6 Policy + 低时延切片承载推理业务，实现业务广覆盖以及用户至任意推理池的三层可达。用户侧设备标识推理业务，网络控制器感知全网状态，在全网为推理业务进行低时延选路，保障推理业务体验。通过端到端切片隔离和数据加密保障业务独立性和数据安全。对算力城域网的关键技术要求如下：

- 确定性时延：推理业务端到端低时延 RTT 小于 10ms，保障用户体验。
- 广覆盖接入：用户接入侧到各级应用推理池 L3 层联接可达，实现 2C/2B/2H 业务的统一接入、融合承载。
- 切片隔离：端到端切片隔离，保障业务间的独立性。
- 数据加密：推理数据加密传输，确保数据安全。

八、总结与展望

本白皮书聚焦算力城域网典型算力业务的需求和特征，对算力城域网应用场景、网络架构、组网方案和技术展开探讨，并积极推动算力城域网的技术研发和现网验证。

算力城域网的规划和建设应由用户算力需求与算网技术发展的双向驱动。中国电信联合华为公司在上海、浙江、广东、安徽等地积极开展算力城域网现网验证，并针对友好用户逐步推进服务试商用；并在业界率先提出了基于算力业务感知的自适应路由技术，在 IETF 孵化 FANTEL 工作组，开启了算力城域网的新协议、新技术探索。

我们希望通过本白皮书的研究与分析，激发更多业界同行的参与和讨论，期盼与众多合作伙伴一起携手并进，汇聚行业力量，共同打造广覆盖、高弹性、广域无损、超高可靠、智能化的算力城域网络。

附录 A：术语与缩略语

中文名称	英文缩写	英文全拼
AI	Artificial Intelligence	人工智能
AIDC	Artificial Intelligence Data Center	面向智算的新一代 数据中心
AIGC	Artificial Intelligence Generated Content	人工智能生成内容
CPE	Customer Premise Equipment	客户终端设备
DC	Data Center	数据中心
ECN	Explicit Congestion Notification	显式拥塞控制
EVPN	Ethernet Virtual Private Network	以太虚拟专网技术
FLOPS	Floating-point Operations per Second	每秒浮点运算次数
GAGR	Gross Annual Growth Rate	年均复合增长率
GB	GigaByte	千兆字节（字节）
Gbps	Giga bits per second	千兆比特每秒
HPC	High Performance Computing	高性能计算
IDC	Internet Data Center	互联网数据中心

IDC	International Data Corporation	国际数据公司
IP	Internet Protocol	网际互连协议
IPv6	Internet Protocol Version 6	网际互连协议第 6 版
LLM	Large Language Model	大语言模型
L3EVPN	Layer Three Ethernet Virtual Private Networks	三层以太虚拟专网
MB	MegaByte	兆字节（字节）
PFC	Priority-based Flow Control	基于优先级的流控
RTT	Priority-based Flow Control	往返时间
QP	Queue Pair	队列对
TB	TeraByte	太字节（字节）
TTFT	Time To First Token	首 token 时间
TPOT	Time Per Output Token	每 token 时间

参考文献

- [1] 以网补算，构筑智算时代新底座[EB/OL](2024-5-50)[2024-8-1]. <https://xueqiu.com>
- [2] Gemini Team, Google. Gemini: A Family of Highly Capable Multimodal Models, 2024.
- [3] Kun Qian, Yongqing Xi, Jiamin Cao et al. Alibaba HPN: A Data Center Network for Large Language Model Training, 2024.
- [4] DeepSeek-AI et al. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning, 2024.
- [5] Patel, Pratyush et al. Splitwise: Efficient Generative LLM Inference Using Phase Splitting, 2024.
- [6] Microsoft. Empowering Azure Storage with RDMA, 2023.
- [7] 云化 IP 城域网中 vBRAS 池化部署研究, 电信科学, 2023
- [8] 百度智能云.智算中心网络架构白皮书[R/OL](2023-6).
- [9] 新质互联网智鉴报告（v1.0）[R/OL](2024-9).
- [10] 算力城域网白皮书[R/OL](2024-11).
- [11] 新质互联网智算数据中心网络建设技术要求[R/OL](2024-11).
- [12] Metropolitan Area Network for the AI Era[R/OL](2025-6).