准印证号：（粤B）L0250077

中兴通讯技术 简讯

ZTE TECHNOLOGIES | 第30卷 第3期 · 2026年3月

视点

专题：智算网络

内部资料 免费交流

扫码体验移动阅读

1996年创办
第30卷 总第449期
2026年03月 第3期（月度出版）

中兴通讯技术（简讯）
ZHONGXING TONGXUN JISHU（JIANXUN）

主管：中兴通讯股份有限公司
主办：中兴通讯技术杂志社
出版：《中兴通讯技术（简讯）》编辑部

编辑部地址：深圳市科技南路55号中兴通讯研发大楼
发行部地址：合肥市金寨路329号国轩凯旋大厦12楼
发行部电话：0551-65533356
https://www.zte.com.cn/china/about/magazine

发行范围：国内业务相关单位
印数：5000本

设计：深圳市奥尔美广告有限公司
印刷：深圳市旺盈彩盒纸品有限公司
印刷日期：2026年03月25日

李强
中兴通讯承载网产品总经理

以网强算，智启未来：构筑智算时代新底座

当前，人工智能正加速重塑数字经济格局，成为驱动产业变革的核心力量。大模型在互联网、政务、金融、制造、交通等领域的规模化落地，催生持续调用、实时响应的新服务模式，“毫秒用算”正成为衡量智能服务响应能力的关键标尺。伴随“东数西算”工程深化推进，全国一体化算力布局加快完善，超大规模智算集群加速部署，跨数据中心的协同训练、分布式推理等多元负载，对网络吞吐率、尾延迟及故障恢复能力提出更高要求。网络作为连接算力资源的“主动脉”，已成为决定算力服务品质的核心引擎。

然而，运营商在智算网络建设中仍面临多重挑战：算力分布碎片化，跨域协同难，易形成“算力孤岛”；传统网络架构难以满足大模型场景下的大带宽、低时延、高可靠传输需求，时延抖动与尾延迟显著影响训练收敛速度与推理服务质量。尤为突出的是，网络与算力资源调度长期割裂，缺乏端到端SLA感知与动态调优能力，制约资源高效协同，推高运维复杂度与成本。如何实现算网深度融合，构建高效、弹性、可调度的智算网络，已成为保障高质量算力服务落地的关键所在。

面对这一系统性需求，中兴通讯秉持“以网强算、以智赋网”理念，打造覆盖数据中心与广域网络的全栈智算解决方案。在Scale-Out场景下，推出GSE框盒智算方案，通过N2N模式，在不更换现网主流智能网卡的基础上为客户提供无损智算体验；依托自研高通量芯片，构建高密度、低功耗、易扩展的智算底座，全面打通“算力最后一公里”。在广域侧，融合SRv6与网元内生智能技术，构建高带宽、低时延、智能容损、高可靠自愈的传输底座，支持端到端切片隔离与动态调优，实现“数据可流动、算力可调度、服务有保障”的新型算力供给模式。

本期专题聚焦高性能智算网络的技术演进与实践探索，系统呈现中兴通讯在算网融合时代的创新路径。展望未来，中兴通讯将持续深化智能中枢能力，赋能高阶智算网络体系建设，与产业同行共绘高效协同、泛在可达的智联图景，为数字中国筑牢根基。

目次

中兴通讯技术（简讯）2026年第3期

构建算力互联底座，助力算网协同高效发展——面向智算业务的IP网络解决方案

人工智能的蓬勃发展离不开底层网络的支撑，广域IP网络已成为AI技术落地的核心基础设施底座。

视点

专题：智算网络

成功故事

中兴通讯荣获GSMA GLOMO三项大奖

当地时间3月4日，2026年世界移动通信大会期间，GSMA全球移动大奖（GLOMO Awards）评选结果正式揭晓。中兴通讯凭借持续的技术创新与深度的行业融合实践，成功斩获“最佳专用网络解决方案奖”“开放网关挑战奖”和“最佳活动营销奖”，其创新成果再次获得全球通信行业的高度认可与权威认定。

中兴通讯、中国电信联合打造的EasyOn 5G-A-RobotNet解决方案，凭借在5G-A专用网络与具身智能融合领域的技术突破、商用实践及产业引领价值，携手机器人头部企业智元机器人（AGIBOT）、卓益得机器人（DroidUp）成功斩获“最佳专用网络解决方案奖”。

由中兴通讯、中国移动集团网络事业部、中国移动杭州研发中心与京东集团联合打造的AI-Powered Open Gateway解决方案，凭借领先的技术创新与标杆实践价值，斩获GSMA全球移动大奖（GLOMO）“开放网关挑战奖”。

中兴通讯联合中国电信和荣泽天韵，凭借“5G-A赋能无线新模式演唱会直播”项目，荣获GLOMO奖之“最佳活动营销奖”，标志着5G-A在演唱会直播行业的成功商用。

中兴通讯与中国电信双提案斩获GSMA Foundry卓越奖三项大奖

2026年3月1日，GSMA Foundry颁奖仪式圆满落幕。中兴通讯携手中国电信打造的两大创新提案脱颖而出，一举斩获Foundry Excellence Awards 2026两大类别奖项及GSMA Foundry Innovation GLOMO奖共三项大奖。其中，“5G-A赋能多机器人智能协同”方案摘得“企业创新与新收入模式”奖，“AI智能总检一体机”提案包揽“跨领域卓越”奖与“创新GLOMO”奖，彰显了中兴通讯在5G-A和AI技术融合行业应用领域的全球领先实力。

中兴通讯携手中国移动联合发布GigaMIMO创新成果

2026年世界移动通信大会期间，中兴通讯联合中国移动重磅发布GigaMIMO创新解决方案及实践成果。该方案以空域资源深度重构为核心，依托超大规模天线能力升级与AI技术深度赋能，实现网络性能突破性跃升，为沉浸式通信、车联网、具身智能等AI原生新业务提供全域高速、稳定可靠的连接支撑，助力5G-A网络容量与体验全面增强，并推动6G技术前瞻布局与产业演进。

中兴通讯与中国电信联合发布DGN创新方案

3月2日，2026年世界移动通信大会上，中兴通讯与中国电信联合发布5G-A DGN（面向差异化连接的生成式网络）创新解决方案。该方案融合内生智能与多用户分布式MIMO技术，可高效保障个人及行业用户的差异化连接体验，为多模态AI业务发展筑牢网络底座。

中国联通联合中兴通讯发布系列创新终端，共启家庭AI新纪元

3月3日，在2026年世界移动通信大会现场，中国联通携手中兴通讯联合发布联通云智产品的系列新品，包括联通云智家庭屏、联通云智自由屏及联通云智PAD等。这些新品深度融合联通元景大模型，标志着双方在智慧家庭产品上的合作迈入新阶段，双方将共同构建以AI为核心的未来家庭生活新范式。

中兴通讯2025年营收1339亿元，算力营收同比增150%，构筑AI端到端全栈竞争力

3月6日，中兴通讯发布2025年度报告。报告期内，公司实现营收1339.0亿元，同比增长10.4%；归母净利润56.2亿元；扣非归母净利润33.7亿元；2025年度拟派发现金分红总额占归母净利润比例35%。

过去一年是中兴通讯将智算确立为长期主航道的关键之年。面对人工智能快速发展的机遇与挑战，公司持续巩固网络产品领先地位，加大在算力、家庭及个人终端领域的投入与创新，逐步构建覆盖“基础设施—平台—应用”的AI全栈技术能力。2025年，战略落地初见成效，公司营业收入重回快速增长轨道，展现出强劲的发展韧性。

报告期内，公司算力业务实现跨越式增长，全年营收同比增长约150%，占整体营收比重达24.6%，其中服务器及存储营收同比增长超200%，数据中心产品营收同比增长50%；家庭及个人终端业务持续增长，合计贡献营收25.3%。与此同时，在行业周期切换与业务结构变化的双重影响下，公司毛利率阶段性承压。公司坚持高强度研发投入，全年研发费用达227.6亿元，营收占比约17.0%，持续构建AI端到端能力矩阵，为长期竞争力夯实基础。

中兴通讯斩获GTI Awards三项大奖

近日，全球TD-LTE倡议组织（GTI）年度大奖正式揭晓，中兴通讯一举斩获移动人工智能应用奖、创新技术突破奖、创新产品与解决方案奖三项大奖。

中国移动湖北分公司、小米、中兴通讯联合打造的“5G-A×AI赋能武汉小米智能家电工厂”项目荣获“移动人工智能应用奖”。

中兴通讯联合中国移动推出的GigaMIMO创新解决方案获得创新技术突破奖。

中兴通讯旗下努比亚与字节跳动旗下豆包合作开发的首款AI原生手机努比亚M153豆包手机助手技术预览版，斩获创新产品与解决方案大奖。

移动互联终端升级“连接+”战略，中兴通讯发布全球首款AI+Wi-Fi 8室内CPE

2026世界移动通信大会现场，中兴通讯终端业务携全场景AI终端亮相，全面展示AI技术与终端生态深度融合的创新成果。此次展会发布了全球首款AI+Wi-Fi 8室内CPE中兴G6 Ultra，以及全球首款毫米波室外5G-A CPE中兴G6 Max两款新品，同时直播神器中兴TopFlow亮相，为全球用户带来更智能、更高效、更普惠的AI终端体验。

中兴通讯推出大容量模块化液冷CDU 应对高密度算力挑战

2026年世界移动通信大会期间，中兴通讯重磅推出模块化冷板式液冷CDU。该产品支持400kW至2MW弹性扩容，可灵活配置2~5个CDU/柜，集成机柜、配电、冷却多重功能，精准匹配高算力散热需求，为客户提供“省成本、提效率、稳运行”的硬核液冷解决方案。

土耳其电信携手中兴通讯完成全球首个C+L全频一体化1.6Tbps现网试验

近日，土耳其电信联合中兴通讯在伊斯坦布尔完成全球首个C+L（12THz）全频一体化1.6Tbps优智全光网现网试验，大幅提升系统容量、减少备件种类，并实现400GE/800GE业务超高速传输，为土耳其发展超宽、智能化全光网络，提升土耳其乃至欧亚数字经济水平奠定坚实基础。

陈志伟
中兴通讯承载网规划总工

赋能先进AI架构，解锁算力潜能
——智算网络演进趋势分析

2025年至今，全球大模型的发展步入成熟发展期，技术叙事愈发宏大：OpenAI推进“百万GPU”战略布局，并开始部署“星门计划”；xAI推出由20万张H100 GPU卡训练的Grok-4模型，采用标准以太网架构，在多项基准测试中表现优异；Google则以TPU v6/v7与自研OCS（全光交换）网络为技术底座，支撑Gemini 3.0体系化演进。我们认为，模型演进的趋势已定——Scaling Law的生命力依然顽强，但算力规模的扩张正逼近网络的通信墙，智算业务从规模至上转向有效算力，百万级集群呼之欲出，网络不再是AI算力的连接配角，而是决定其上限与效率的中枢神经系统，这迫使机间网络（Scale-Out）标准加速收敛，并引爆机内互联（Scale-Up）的百花齐放。

Scale-Out：技术标准归于收敛，GSE与UEC各领风骚

在机间网络（Scale-Out）领域，核心诉求始终是大规模互联、极致带宽利用率以及确定性的低延迟。2025年以来，随着以太网技术的飞速演进，RoCE（RDMA over converged ethernet）在很大程度上已经开始替代传统的IB（InfiniBand），尤其是在追求成本效益和开放性的互联网领域。RoCE的持续演进，是为了在性能上进一步超越IB。

UEC 1.0确立下一代以太网传输范式

海外由超级以太网联盟（UEC）主导的UEC规范在2025年6月发布了1.0正式版。UEC不仅是对以

太网的修补，而是从物理层到传输层（UET协议）的彻底重构。其核心创新在于：链路层，LLR（link layer retransmission）与CBFC（credit-based flow control，基于信用的流控）协同实现“准无损”；传输层，通过报文级喷洒（packet spraying）技术，将智算网络的整体带宽利用率提升至接近100%。

传统的以太网丢包依赖传输层进行端到端重传，这在超大规模集群中会导致极高的尾部延迟，LLR实现了逐跳（hop-by-hop）的本地修复机制。当交换机检测到链路抖动导致的微小丢包时，直接在物理链路层完成重传，无需触发全局重传。这种近场恢复能力将丢包对模型训练的影响降到了物理级最低，是支撑百万卡规模运行的基石。此外，传统RoCE依赖的PFC（优先流控）机制容易引发死锁或PFC风暴，CBFC采用主动式信用额度管理，发送方必须获得接收方的信用额度授权才能发送数据。这种机制从源头避免了交换机缓存溢出，实现了真正的确定性转发。

在传输层，UEC彻底打破了传统以太网基于流的ECMP负载均衡限制，通过报文级喷洒技术，UET协议允许将同一个AI训练任务的海量数据切分为细粒度报文，均匀分布到网络拓扑中的所有可用路径上，将智算网络的整体带宽利用率提升至接近100%。UET同时支持选择性重传（selective retransmission），网络仅需补发真正丢失的报文，而非回退N个报文全部重传（go-back-N），极大节省了带宽资源，缩短了任务完成时间（JCT）。

目前，最新发布的102.4T交换芯片，包括博通（Broadcom）的Tomahawk 6、美满科技（marvell）的T100，英伟达（NVIDIA）的spectrum6，都开始支持UEC协议。UEC协议下一步的规模部署，标志着以太网在AI领域完成了对IB协议的全面超越。

GSE：中国的UEC，以N2N模式彰显独特优势

国内方面，由中国移动等单位牵头定义的GSE（Global Scheduling Ethernet）已成为中国智算网络的标杆性标准。GSE在技术理念上与UEC志同道合，但其在工程实现上更具中国智慧。

N2N（network-to-network）模式是GSE相比UEC最显著的技术分水岭。UEC的诸多特性（如UET协议）高度依赖于网卡的同步升级，这意味着用户必须采购全新的、支持UEC标准的网卡。而GSE主推N2N模式，核心技术创新主要在网络侧实现。它通过交换机侧的全局资源感知和报文动态切片，兼容现有的标准RoCE v2网卡。同时GSE创新容器（container）技术，基于容器而非报文的喷洒技术，同时满足了流量喷洒的均衡性和降低报文乱序的概率。这种不依赖网卡升级的特性，极大降低了旧有集群的升级门槛，支持异构网卡环境下的统一无损转发，尤其适合中国的部署现状。

2026年下半年，预计支持GSE N2N的51.2T国产化芯片将会发布，标志着国内在智算Scale-Out领域真正有了原创性的产品。

Scale-Up：百花齐放，技术路线尚未收敛

在机内/机柜内扩展（Scale-Up）领域，2026年虽然英伟达凭借NVLink 6.0及其闭环生态在性能上依然领先，但海外以UALink和SUE/ESUN为代表的开放阵营正通过不同的技术路径实现快速超车。

2025年，AMD、Intel、Google等巨头联合发布UALink 1.0规范。该规范保留了开发者熟悉的总线编程模式，支持内存一致性协议，使得GPU之间可以像访问本地内存一样互相读写。通过引入以太网PHY，UALink突破了传统PCIe的距离限制，支持多达1024个加速器组成的超大规模Fabric域。支持UALink的芯片预计2026年发布。

博通（Broadcom）则代表了另一条路径——网络型路线。博通推出的SUE（Scale-Up Ethernet）、OCP发起的ESUN（Ethernet for

Scale-Up Networking），主张利用以太网生态的极致成熟度来实现Scale-Up。方案去除了复杂的IP层，修改了MAC层部分逻辑，通过简化报文头和报文转发逻辑，实现极致低延迟。博通Tomahawk6已经兼容支持部分SUE特性，支持全部SUE特性的TF2预计2026年发布。

与Scale-Out已基本收敛至以太网不同，机内互联在总线型与网络型之间尚未达成最终统一，Scale-Up目前仍处于百花齐放的战国时代。这种技术路线的博弈，预示着未来2—3年内，谁能率先提供比肩NVLink性能且具备开放生态的互联方案，谁就将获得最后的胜利。

底层硬件技术的快速发展

为了支撑百万卡集群的宏伟蓝图，底层硬件技术也在经历翻天覆地的变化。

112G SerDes规模部署，224G SerDes 2026年商用部署，448G SerDes已在路上

SerDes是定义智算带宽的元技术之一。在Scale-Out侧，112G SerDes海外和国内方案已成熟，并实现大规模部署。224G SerDes海外头部芯片厂商推出新一代芯片，国内厂商在积极跟进，我们判断国内在2年内会有自研224G SerDes芯片推出。同时，448G SerDes的开发已经开始。448G SerDes在封装、信号完整性、散热上会有很多新的技术挑战，预计英伟达用于Scale-Up的nvswitch6会是首个采用448G SerDes的芯片。

电的演进，从PCB到NPC和CPC

随着SerDes速率从112Gbps迈向224Gbps甚至448Gbps，信号频率的提升使传统PCB走线面临严峻的插损挑战。针对该挑战有NPC（near package cable）和CPC（co-packaged copper）两种方案。NPC方案引入Flyover Cable技术，通过芯片—PCB NPC插座—电缆—光模块的路径，利用电缆替代长距离PCB走线，显著优化信号质量。CPC芯片封装直接出电缆，信号直接从芯片封装经Flyover电缆传至模块。我们判断224G时代NPC是一个更好的选择，CPC还处于探索阶段。

光的演进：LPO是可选路径，CPO是必由之路，NPO是解耦选择

光模块功耗已占到网络总功耗的50%以上。LPO（线性驱动可插拔光模块）在2025年下半年经历了从质疑到规模应用的过程，其通过省去DSP显著降低了时延和功耗。然而，面对1.6T及更高速率，CPO（共封装光学）会是终极解决方案，已在部分头部客户开始小规模部署。NPO（近封装光）可以把switch芯片/GPU芯片和光引擎（optical engine）解耦，得到了部分国内互联网和GPU厂商的看好，在国内开始试点。预计CPO会在102.4T时代得到更多的部署，在204.8T时代成为主流部署方案。

智算网络，开放与自研的共振

中兴通讯认为，AI大模型时代的智算网络正处于前所未有的剧变期。我们坚持开放解耦与深耕底层的双轮驱动战略。在Scale-Out领域，我们积极拥抱收敛趋势，作为GSE的核心参与者，中兴通讯已经推出支持GSE的框盒智算方案，通过N2N模式为客户提供无需更换网卡即可实现的无损智算体验。在Scale-Up领域，我们参与百花齐放的竞争，通过牵头Clink并深耕Onlink，致力于打破私有协议垄断，为国产GPU和自研加速器构筑高品质的超节点连接。

智算网络的演进是一场长跑，中兴通讯将继续秉承开放合作的姿态，与全球及国内产业界伙伴一同，在百万卡时代的算力洪流中，架设起坚实、高效的信息通途。ZTE中兴

陶文强
中兴通讯承载网规划总工

高性能智算广域网络演进趋势及关键技术

人工智能应用正加速从单一大模型向智能体（AI Agent）体系演进，AI发展正式迈入“智能体元年”。与传统响应式AI工具不同，AI Agent具备持续交互、状态记忆、环境感知与自主决策等类人能力，正推动人机交互范式从“功能连接”向“持续服务连接”深刻转型。据IDC预测，到2030年，全球AI Agent数量将快速攀升至22.16亿，成为连接用户、服务与物理世界的数字基础设施，重塑终端形态、服务模式与产业生态格局。

在此背景下，AI流量特征发生根本性转变，AI Agent的7×24小时在线特性催生“低基线稳态、间歇性突发、弹性峰值”的新型流量模型。测算表明，AI驱动的网络流量年复合增长率高达51%，预计至2033年，整体流量规模将增长5至9倍。其中，东西向流量占比显著提升，逐步超越传统南北向流量，成为广域网络承载的核心挑战。

与此同时，AI算力集群正从万卡级向十万卡、百万卡规模快速演进。超大规模并行训练、云边端协同推理等新模式对网络提出前所未有的综合性能要求。传统“尽力而为”的网络架构已难以支撑AI业务对高吞吐、低时延、低抖动与高可靠性的承载需求。

面向未来，广域网亟需从“通用连接管道”向“算力使能网络”转型，为AI发展提供关键支撑。

智算业务场景和网络承载需求分析

智算业务的多样化催生了高度异构的流量图

▼表1 智算业务分析

场景	面向用户	场景特点	传输频次	传递数据类型	智算网关角色	流量特征
样本入算	ToB	海量样本入算、数据快递	高：每用户数天一次	样本数据、用户数据	同时入算的多用户样本数据、用户数据汇聚	大量数据持续上传
存算分离	ToB	样本数据随训随取，样本数据不落盘	低：每用户数月一次	样本数据	同时进行训练的多用户样本数据汇聚	每轮轮迭代数据上行传递一次
PD分离	ToB/ToC	推理预填充（P）与解码（D）拉远	高：推理会话频发	KV Cache	P池向D池推送数据	每轮推理：1份KV Cache传递
分布式推理	ToB/ToC	首末层在本地，中间层在远端；输入输出不出园区	高：推理会话频发	KV Cache、前向激活值	多用户KV Cache与激活值汇聚	每轮推理：1份KV Cache+2份激活值
边云协同后训练	ToB	首末层在本地，中间层在远端；训练数据不出园区	低：每用户数月一次	后训练参数面数据	多用户参数汇聚与下发	每轮迭代参数面数据上下行各传递一次
分布式训练	ToB	多域算力协同并行，模型与数据全域高效训练	单次持续时间长	训练参数面数据	参数面数据传递	周期性大流量参数面数据

谱。智算业务典型场景涵盖样本入算、存算分离、PD分离、分布式推理、边云协同后训练及分布式训练六大类型，这些场景在数据规模、传输频率、时延敏感度和突发性等方面差异显著，并呈现“规模跨度大、时延敏感度高、周期性与突发性并存”的共性特征（见表1）。

其中，PD分离与分布式推理中频繁传输的KV Cache具有小包高频、时延敏感的特点，要求网络具备微秒级响应与低抖动保障；而分布式训练的参数同步则需持续、无损、高吞吐的TB级数据传输；样本入算涉及PB级冷数据广域迁移，强调带宽利用率与时效性；存算分离场景单次数据量大，对网络带宽和传输稳定性要求高。

面对差异化、高动态的业务需求，网络亟需向“业务感知、按需切片、弹性适配”的智能体系演进：基于“网业分离”设计理念，构建面向多业务场景的统一承载平台；通过为不同业务类型设置独立的业务锚点，实现各业务平面的灵活独立演进与按需调优，既保障关键任务的服务质量，又有效隔离业务间相互影响，显著提升网络的可维护性与扩展性。

此外，端到端无损是关键场景的保障基石。RDMA对丢包极为敏感，即使2%的丢包也可能导致吞吐量断崖式下降。因此，在万卡级训练、实时推理等高同步强度任务中，必须构建跨广域无损能力，涵盖显式拥塞控制、动态速率调节与微突发缓冲机制，确保算力高效协同、零空转。同时，部分非实时场景如样本上传、参数异步更新等对时延与丢包容忍度较高，可引入智能有损机制，在保障整体服务质量及控制成本前提下提升带宽复用效率，形成分级承载策略。

综上，高性能智算广域网需构建“任务驱动、差异服务、端网协同”的新型架构，在统一承载基础上支持灵活演进，推动网络从“算网融合”向“网智共生共融”持续演进。

高性能智算广域网架构及关键技术

高性能智算广域网需支撑多样异构业务，通过端、网、管协同实现业务感知与资源动态匹配，构建弹性、无损、智能的承载体系，为智算任务提供可保障的传输服务。

端网协同架构

高性能智算广域网通过和端侧、算侧深度融合，进一步提升端到端整体性能。如图1所示，高性能智算广域网整体架构主要由端、网、管三类核心元素构成：

端侧，主要指智算网关，端侧作为数据传输的发起端，需具备动态速率调节能力，通过端网协同机制，端侧可接收网侧传输建议信息，优化传输速率，实现与网络状态的动态匹配。

网侧，包含广域网PE设备和转发设备，在保障公平性的情况下提供尽可能高的吞吐。通过与端侧的协同，网络能够感知业务流量特征，实施主动拥塞避免，提升传输效率与稳定性。

管控，主要由网络控制器和统一调度平台组成，负责广域网络流量控制和资源调度。通过与端侧、网侧的协同，控制器可基于传输任务特性和网络资源状况，动态分配带宽等资源，优化多任务并发下的资源利用率，提升整体网络效能。

端侧系统通常部署于智算中心或数据站点，端侧与网络通过端网协同进行拥塞控制，交互智算任务的流量需求及网络能力，协商传输速率。网络据此可提供资源规划及调度支持，保障各类任务的传输需求及资源承诺。

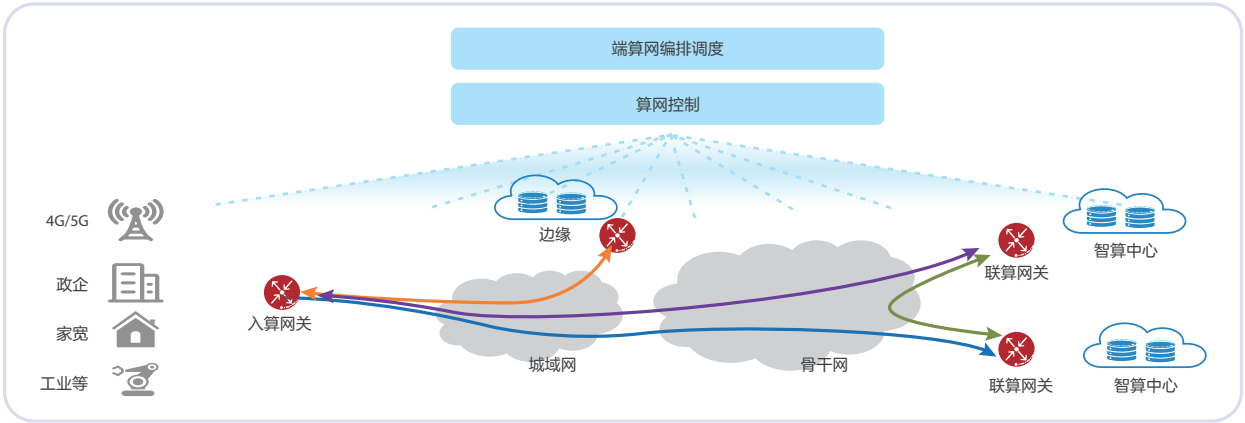
其中广域网可通过确定性无损技术提供高保障服务，网络预留带宽队列及buffer等资源并实施入口准入控制等手段，实现网络无拥塞、无丢包及可承诺时延。对于非关键或容忍丢包的业务，网络也可以通过智能QoS和智能ECMP等技术提供弹性承载：智能QoS可优化流量突发导致的丢包，实现吞吐最大化与最低时延的目标；智能ECMP则通过感知流量的生命周期及流量大小分布等动态特性，突破传统静态Hash机制的限制，实现多路径更均衡的负载分担。

端网协同关键技术

端网协同关键技术聚焦提升传输效率与网络利用率，通过拥塞控制、智能调度与路径优化等手段，实现流量与网络资源的动态适配，保障多场景下高性能传输需求。

- 智算网关与网PE协同的拥塞控制技术

智算场景中，突发流量在广域网传输时可能引发瞬时拥塞、丢包和排队延迟。传统拥塞控制机制下，端侧缺乏对网络的带宽资源的量化感知，导致速率调整不平滑，收敛速度慢，进而影响吞吐。为此，高性能智算广域网引入端网协同拥塞控制机制，基于任务级应用需求与端侧协商发送速率，通过向端侧动态分配和授权流量发送配额，预防网络拥塞，同时实现基于配额的资源调度、准入控制和流量调控，满足最优完成时间目标。通过端网协同，客户端和服务端可更精细、高效地调节发送速率。网络由此增强对流量的调控和资源调度能力，保障各类任务的传输需求及资源供给，提供可预测的网络行为，有效缓



▲图1 高性能智算广域网端网协同总体架构

解网络拥塞。

- 智能大象流识别和报文级喷洒（packet spraying）技术

智算场景中，大流量传输常因广域网资源不足或分配不均导致链路利用率下降；多流并发传输海量数据时，易引发网络资源过载，造成拥塞和吞吐下降。为提升带宽利用率，需实现多路径负载平衡，以达成低延迟、零丢包和高吞吐量的传输性能。传统静态ECMP负载均衡的核心问题在于固定哈希策略无法适应动态流量特征，易导致流量分配不均（如大象流独占路径）、协议层冲突（如TCP报文乱序），在AI训练等高并发、高突发性场景下问题尤为突出。

为此，高性能智算广域网引入报文级喷洒技术，将同一个AI训练任务的数据流切分为细粒度报文，并均匀调度到网络所有可用路径上，充分复用多路径带宽资源，显著提升整体带宽利用率。

线卡集成AI-flow分析组件，实时识别大象流并监控其生命周期及带宽等基线统计信息：

- 流量带宽度量：基于端口负载、流带宽、队列深度等因子，识别潜在拥塞风险，为路径选择提供依据；
- 生命周期度量：监测流的生命周期与实时统计数据，动态选择进入或者离开策略。

基于AI-flow的流量基线信息，系统综合评估流的传输特征与各ECMP路径的负载状态，动态计算最优路径。通过多路径流量智能调度，实现带宽聚合、毫秒级故障切换，有效突破传统静态哈希机制的性能瓶颈。

标准进展

为应对智算场景对低时延、高吞吐、低丢包的严苛需求，互联网工程任务组（IETF）已在传输、管控与路由等多个技术领域启动相关标准研究与讨论。在传输层面，SCONE、TSVWG、CCWG等工作组聚焦于RDMA（包括RoCE）与TCP、QUIC等协议的适配机制，并深入研究CUBIC、BBR等拥塞控制算法的优化方案。同时，IETF正在推进“高性能广域网”（HP-WAN）新课题的研究与讨论，明确广域网需支持高速、低延迟与超高容量的应用场景，满足高吞吐与低时延并重的基础能力需求。

国际电信联盟ITU SG13在算网架构、资源建模、业务增强方面提供了顶层设计和参考架构，为全球算网融合标准化进程奠定基础。国内以CCSA为核心，已启动《面向智算业务的广域网总体技术要求》等行业标准立项和制定工作，围绕广域无损、应用感知、确定性承载等关键方向初步形成体系化布局。

构建面向AI未来的智算广域网络新格局

中兴通讯认为，人工智能已进入新一轮高速发展周期，尽管传统规模定律（Scaling Law）的边际效益正面临挑战，但业界对算力规模的需求仍在持续增长，驱动重心正从大规模预训练向后训练、推理及多智能体协同等多元化场景延伸。在此背景下，AI集群规模从万卡向十万卡乃至百万卡演进，已成为行业发展的关键趋势。作为AI基础设施的核心组成部分，广域网络亟需同步升级，迎来新一轮技术革新。

在面向AI业务的广域承载网络的演进中，无损与有损技术并非优劣对立，而是面向不同业务需求、部署条件与规模边界形成的差异化技术路径。二者并非替代关系，而是长期共存、按需分层、智能适配，最终形成“关键任务无损保确定性、非关键流量有损保普惠性”的智算承载新格局。

当前多种技术路径并行发展，创新活跃，尚未形成统一标准，正处于关键技术路线的探索与收敛期。网络已成为驱动算力释放的关键使能者，唯有通过端网深度协同与技术架构持续创新，方能构建面向超大规模AI集群的高效、弹性、智能、确定性广域网络，为人工智能的下一阶段发展提供坚实支撑。ZTE中兴

构建算力互联底座，助力算网协同高效发展

——面向智算业务的IP网络解决方案

中兴通讯 冯志坚，庄严

图片由AI辅助生成

冯志坚
中兴通讯BN产品创新方案总工

庄严
中兴通讯BN产品MKT及方案团队副部长

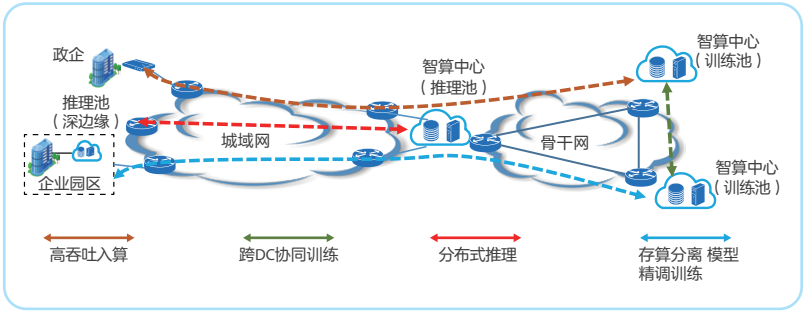
中国围绕人工智能发展已构建了全链条、多层次的政策促进体系，2025年国务院出台《关于深入实施“人工智能+”行动的意见》，进一步提出更具体的阶段目标，要求实现人工智能与重点领域深度融合。人工智能的蓬勃发展离不开底层网络的支撑，广域IP网络已成为AI技术落地的核心基础设施底座。面对AI大模型的分布式训练、跨地域数据传输、存算拉远等场景对网络带宽、低时延、高可靠性提出的新要求，运营商IP网络通过技术升级与方案创新，为AI发展提供关键保障。

智算业务场景的网络要求及关键技术

在AI大模型发展驱动下，IP网络面临高吞吐入算、跨DC联合训练、跨DC分布式推理、跨DC存算分离等智算业务新场景（见图1）。各场景需要依托高速传输、智能调度等关键技术，适配跨地域算力协同需求，破解数据传输、算力整合、安全合规等难题，支撑一体化算力网络高效运转。

高吞吐入算

“东数西算”战略推进下，形成“东部供数、西部算力承接”的核心格局，高吞吐数据入算需求显著。AI大模型全生命周期训练依赖海量样本入算，其中预训练数据量达PB级，后训练与微调阶段随用户规模扩张，整体数据量呈激增态势。此类数据以周期性批量传输为主，需依托国家算力枢纽节点布局，实现跨区域、跨资源池高效流转，为“东算西训”“东数西存”核心场景提供支撑。



▲图1 智算关键业务场景

高吞吐入算场景的网络要求及关键技术如下：

- 弹性高吞吐：支持100Mbps~100Gbps带宽分钟级开通、秒级动态调整，结合400G/800G全光网络技术，保障TB/PB级海量数据稳定高速传输。
- 智能调度均衡：基于SRv6技术构建动态路由体系，通过全局与网络级双重负载均衡，实现数据流智能拆分与多路径并行传输，优化跨枢纽节点传输路径，降低端到端时延。
- 租户隔离安全：采用层次化网络切片技术，隔离样本入算流量与普通业务流量，满足多主体租户级安全隔离需求，保障数据传输合规性。
- 便捷资源联动：支持企业单点接入直达通算、智算、超算等异构资源池，适配全国一体化算力网“枢纽－集群－节点”三级架构，实现跨池数据按需调度。
- 差异化服务计费：自动识别业务优先级并分配专属弹性传输管道，提供多维度计费模式，契合算网协同运营的市场化需求。

跨DC联合训练

跨DC联合训练的核心是解决单一数据中心资源短缺、算力碎片化问题，通过联动多地数据中心实现算力池化整合，支撑万亿级参数大模型等超大规模训练任务落地。训练过程中，每轮迭代产生的中间数据达TB级，且数据同步依赖对丢包高度敏感的RDMA协议；同时，训练产生的“大象流”易引发负载不均衡，仅0.1%丢包即会显著降低算效，对网络传输要求严苛。

跨DC联合训练的网络要求及关键技术如下：

- 大带宽支撑：需部署400G及以上传输链路，高端场景升级至800G广域无损传输，构建跨地域算力协同的高速互联底座。
- 广域负载均衡：对整网流量实施统一规划，实现数据中心网络与广域网一体化调度，通过全局路径优化避免局部链路过载，保障网

络高吞吐性能。

- 高收敛比组网：结合集合通信算法与网络优化技术，按最优收敛比组网，平衡算效与建网成本，降低基础设施部署投入。

跨DC分布式推理

跨DC分布式推理通过模型拆分或并行化部署至跨地域节点，行业主流采用PD分离框架优化效能，将计算密集型Prefill阶段与访存密集型Decode阶段解耦，大幅降低推理成本。边缘节点执行Prefill计算，中心节点承接Decode任务并留存对话上下文及高热度KV Cache，缩短响应链路，实现低时延、高稳定的推理服务交付。

跨DC分布式推理场景的网络要求及关键技术如下：

- 无损传输：依托RDMA（如RoCEv2）构建无损传输环境，搭配PFC流量控制与ECN拥塞通知，实现微秒级时延与近零丢包。
- 云边协同保障：借助SRv6优化转发路径，结合RDMA加速KV Cache的预取与同步效率。

跨DC存算分离

在通用大模型的基础上，用行业专属数据进行模型精调，可以采用跨DC“存算分离”方案，该方案平衡企业敏感数据保护与高效用算需求。通过数据存储与训练算力地理分离，本地留存核心敏感数据规避跨域风险，远端智算中心承载训练任务，破解“数据不出域”与“高效用算”的核心矛盾。

跨DC存算分离场景的网络要求及关键技术如下：

- 高性能传输：核心链路部署100G及以上速率，骨干链路升级至400G级别，适配高频大批量“大象流”传输。
- 智能调度适配：通过控制器实现整网编排，自动调度动态规划传输路径；结合SRv6切片划分专属通道，弹性调整带宽适配算力需求。
- 安全隔离防护：依托SRv6实现租户隔离；启用零信任认证严格校验节点身份，保障数据合规与隐私安全。
- 高可靠保障：链路层面采用主备冗余，跨地域搭配专线与VPN备份；通过BFD毫秒级故障检测配合秒级重路由，部署监控与故障分析系统，保障业务持续运行。

智算IP网络整体解决方案

智算业务涵盖“东数西算”、模型训练、AI推理、存算分离等多元场景，IP网络解决方案需围绕高带宽、低时延、无损传输、安全隔离、智能调度及高可靠核心诉求，通过架构优化、技术适配、策略配置等多维度设计，全面匹配业务差异化需求。

组网架构优化

通过高速全互联、中心边缘协同、全域跨域调度，构建高效、弹性、广覆盖的网络架构。

- 核心层与接入层设计：采用高聚合度全互联核心架构，核心层部署400G/800G传输能力，接入层按节点类型灵活配置400G/100G接入，实现算、存节点与核心层高效互联。
- 中心-边缘协同设计：构建“骨干网+边缘节点”分布式架构，边缘节点就近接入用户网络，缩短响应链路。
- 跨区域互联设计：匹配全国一体化算力网“枢纽-集群-节点”三级架构，搭建跨区域互联通道，支持企业单点接入直达各类异构资源池，实现跨池数据按需调度与算力协同。

传输性能提升

通过高速大带宽与弹性调度支撑海高效传输，同时依托无损低时延技术提升算力利用率。

- 带宽弹性扩容：核心链路部署100G及以上速率，跨地域骨干链路升级至400G/800G，满足TB/PB级数据批量传输与高频“大象流”交互需求；部署弹性带宽调度系统，支持100Mbps~100Gbps分钟级开通、秒级调

整，适配潮汐式流量特征。

- 无损低时延优化：全面部署RDMA（如RoCEv2），依托“零拷贝”特性将传输时延压缩至微秒级，搭配PFC与ECN机制构建广域无损环境，严控丢包率，算效劣化不超过5%。

智能调度与负载均衡

依托智能路由与精细化流量管理，实现全网资源动态优化、业务高效稳定运行。

- 全局动态调度：基于SRv6可编程路由构建动态路由体系，结合控制器实现整网智能编排，通过优化算法动态规划路径，实现数据流拆分与多路径并行传输。
- 流量适配：为大小流分配不同优先级，启用优先级队列调度；部署自动调度机制，根据业务SLA动态调整资源分配，最大化带宽利用率并应对流量突发。

IP网络切片

基于SRv6构建差异化网络切片，为不同AI业务提供专属SLA保障，并通过弹性调度与可视化监控，实现资源智能适配与业务稳定运行。

- 切片体系构建：基于SRv6构建网络切片体系，按业务类型划分批量传输、大模型训练等专属切片，每个切片配备独立带宽、时延、安全SLA保障。
- 切片资源弹性调度：通过切片控制器，实时感知切片内流量变化与资源占用，动态分配资源。
- 监控保障：部署切片可视化监控系统，实时采集时延、丢包、带宽利用率等指标，支持故障快速定位与自愈，保障切片内业务稳定。

安全隔离与合规保障

采用层次化网络隔离与全链路加密防护，实现智算业务安全隔离与数据合规传输，保障多租户环境下业务稳定与数据安全。

- 层次化隔离：以IP网络切片为核心，结合VLAN/VRF技术为不同租户、业务划分独立传输通道，深度隔离智算流量与普通业务流量，防止故障或拥塞扩散。
- 全链路安全防护：加密防护贯穿传输全过程，可以引入IPsec加密、零信任认证、节点身份校验，确保敏感数据安全。

高可靠与智能运维

通过毫秒级故障倒换与多重冗余机制保障业务高可靠运行，同时依托可视化智能运维与差异化计费体系，支撑算网融合的高效运营。

- 多重冗余保障：部署BFD毫秒级故障检测配合50毫秒极速倒换，保障跨域训练、推理等长周期任务连续运行。
- 智能运维与计费：部署随流检测等可视化监控工具，实时采集核心指标。建立差异化服务与计费体系，提供多维度计费模式，契合算网协同市场化运营需求。

总结与展望

智算IP网络作为AI技术规模化落地的核心基础设施，针对高吞吐入算、跨DC联合训练等多元业务场景，通过组网架构优化、传输性能提升、智能调度、安全隔离等多维度方案，满足AI业务对高带宽、低时延、高可靠的核心诉求。

未来，随着智能体互联网的加速演进，智算IP网络将成为智能体间数据交互、协同决策的核心枢纽，推动AI应用从单点智能迈向全域协同智能。同时，AI路由器的持续创新将为智算网络注入更强动力，实现流量的精准识别与智能调度，提升网络吞吐率与运维效率，通过数据自学习不断挖掘行业价值，助力算网协同向更智能、高效的方向发展。

面向未来，智算IP网络将持续深化与AI技术的融合创新，构建数字经济时代的算力互联底座，为人工智能产业高质量发展提供有力保障。ZTE中兴

智能体互联网（IoA）构建：核心技术与网络演进

吉晓威
中兴通讯IP网络规划专家

2025年，中国AI应用市场延续了“高位倍增，强劲渗透”的态势，月活跃用户规模从2024年的1.35亿增长到5.44亿，一年内翻了两番。以OpenClaw为代表的自主智能体兴起标志着AI应用逐渐从“单点工具”向自主协作的“社会化群体”演进。这种演进本质上是技术、需求与进化规律共同驱动的结果：技术层面实现了从next-token-prediction（下一词预测）到next-state-achievement（下一状态达成）的跨越。前者决定了对话式AI仅能聚焦单一文本生成、简单问答等闭环任务，受限于领域精度不足、缺乏自主任务调度与状态适配能力，无法应对跨环节、跨模态的复杂场景；而随着AI价值向产业级流程化应用落地，市场对跨节点协同、高容错性的需求，倒逼技术向next-state-achievement升级。依托该模式，可实现多智能体间的状态感知、分工适配与数据互通，再叠加大模型的任务规划能力、协作算法及分布式算力架构的支撑，真正具备复杂协同能力。

传统互联网/物联网（Internet/IoT）本质上是由人类驱动的“信息互联网”，其核心功能是实现人、机、物之间信息的连接与传递；而随着AI应用的爆发式发展，智能体不再是孤立的工具，而是智能数字社会的“公民”，智能体互联网（internet of agents，IoA）应运而生，智能体间的协作，也将开启前所未有的、高度动态且语义丰富的通信流量新模式。

智能体互联网的挑战与核心功能

构建一个开放、可扩展的智能体互联网，首先需要为其建立一套基础性的“社会运行规则”。这套规则需要解决几个根本性问题：在一个拥有海量、异构智能体的世界里，如何唯一地标识每一个智能体？如何让一个智能体能够找到另一个具备所需能力的智能体？如何确保他们之间的交互是高效、可信且安全的？如何让他们能够跨越不同平台和框架的壁垒，理解彼此的意图并高效协同完成任务？

这些挑战催生了IoA的核心技术架构，其核心能力可以概括为四个层次：智能体标识与发现层、智能体认证与安全层、智能体交互与协作层，以及底层承载网络。前三层构成了智能体交互的Overlay层，而最后一层则是所有这些交互得以发生的物理基础。当前针对如何实现上述功能，国际和国内的产业界学术界提出了两种差异显著的技术哲学和实践路径，分别以AGNTCY框架和智能体网关为代表。

本文聚焦于与承载网络关联密切的“发现”与“路由”问题的解决方案，分析两种框架下的实现机制，并分析由此引发的网络从“连接管道”向“智能协作承载平台”的演进中的关键技术要求。

AGNTCY框架

AGNTCY是由Linux基金会托管，得到思科、谷歌云、戴尔、红帽等科技巨头支持的开源项目，其目标是成为智能体互联网的“TCP/IP套件”。其核心思想是在现有的TCP/IP网络协议栈之上，构建一个专属于智能体的、标准化的应用层基础设施，其构成组件如图1所示。

智能体间的发现主要通过“智能体目录服务”（agent directory service）的中心化或分布式应用服务来实现。每个智能体启动时，会向这个目录服务器注册自己的唯一标识、功能描述以及当前所在的网络位置（如一个API端点域名）。ADS维护分布式哈希表（DHT，distributed hash table），用于高效查找与检索目录记录，DHT将智能体技能映射到记录标识符，从而可根据能力快速发现相关智能体。智能体间的消息路由由SLIM（secure low-latency interactive messaging）实现，通信客户端（智能体）定义采用4段式结构：organization/namespace/service/client，由SLIM节点（SLIM node）从ADS同步到DHT后，实现智能体间的点到点E2E消息和群组E2EE消息转发。

智能体网关

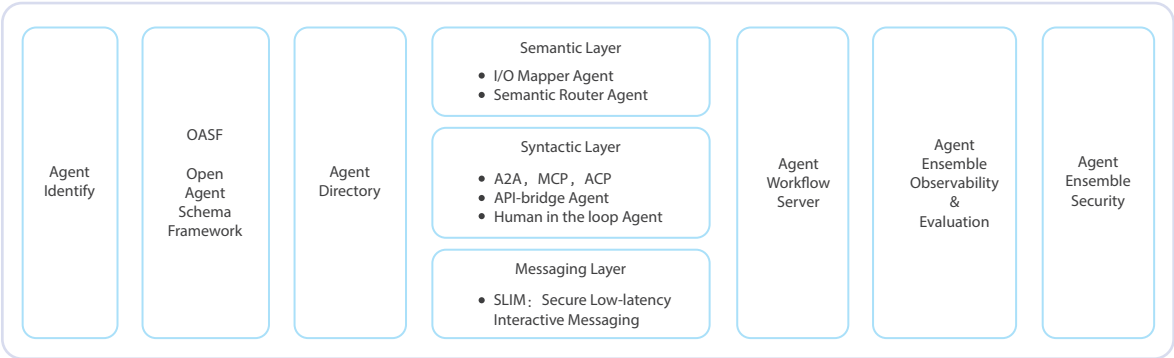
智能体网关基于承载网络自身进行深度革新，在国内受到业界广泛关注。其核心理念是：为了满足智能体间通信对确定性、安全性和效率的极致要求，需要在网络的关键位置（如城域网边缘、数据中心出口）部署专用的智能体网关。智能体网关不是简单的应用服务器，而是一个深度融合了网络转发、身份认证、语义解析和策略执行能力的智能节点。

智能体网关改变了传统的网络寻址范式。在传统IP网络中，路由基于IP地址（设备位置），而智能体网关致力于实现基于“智能体身份标识”的寻址，这意味着一个智能体在通信时，只需指定目标智能体的身份ID标识，而无需关心其当前运行在哪台服务器、IP地址是什么。智能体网关负责维护身份与位置的动态映射，并完成报文的转发。

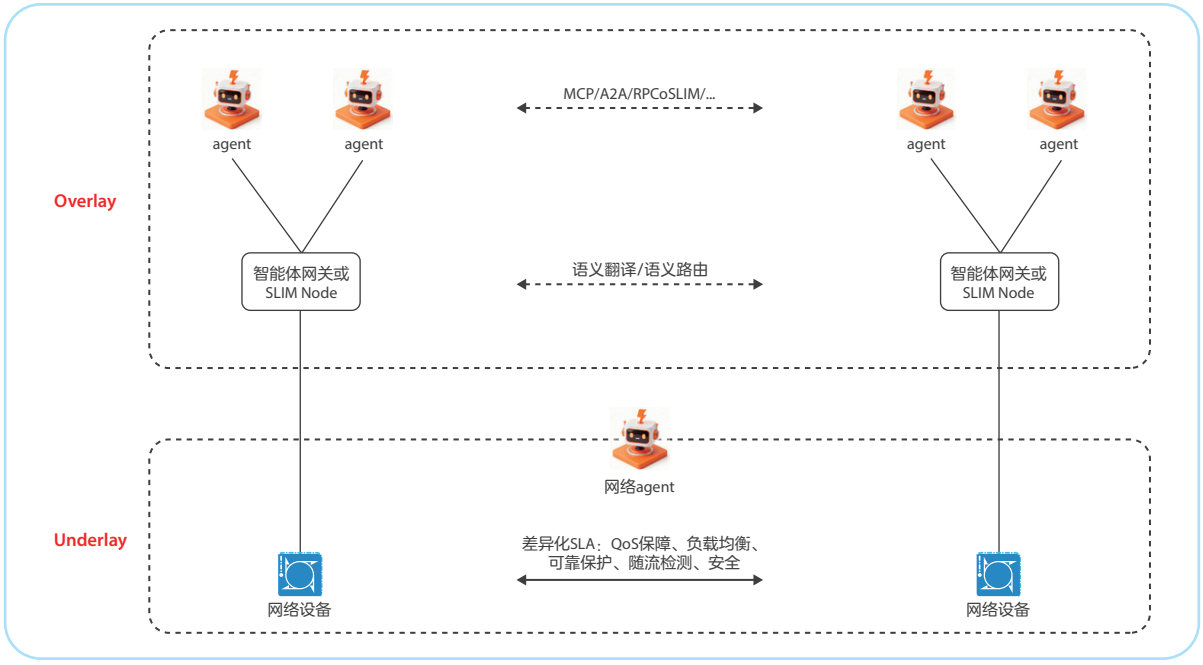
更进一步，智能体网关追求在网络层实现语义感知路由：深度解析或关联智能体交互流中的元数据，理解当前通信的任务类型（实时对话、文件传输或视频生成），并基于这些语义信息以及网络实时状态做出更优的路由决策。比如确保两个智能体间的实时对话流量被调度到低时延路径上，而大文件传输任务则被引导至高带宽链路。这种机制将网络提供的“尽力而为”服务提升为“语义驱动”的差异化SLA保障。

智能体互联网框架及网络的新需求

结合AGNTCY与智能体网关两种设计思想，智能体互联网的框架可以提炼为图2。



▲ 图1 AGNTCY组件框架



▲图2 智能体互联网框架

AGNTCY与智能体网关两种设计思路，都是在Overlay网络层设置节点完成语义翻译和语义路由，但在实现细节上存在很大区别：AGNTCY框架下各智能体需要将相对成熟的智能体通信协议（模型上文协议MCP、智能体到智能体协议A2A等）用SDK转换为SLIM消息，由SLIM Node根据消息头中的通道（channel）信息完成E2E或E2EE转发；智能体网关则基于原生MCP、A2A协议，由智能体网关完成语义翻译和语义路由，但对智能体自动发现、技能（skills）的统一定义等方面还待完善。无论哪种路径最终成为主流，多智能体的协作互联都对网络提出了清晰的新需求：

首先，网络需要从“位置寻址”向“身份寻址”演进。以IP地址为核心的寻址体系无法适应智能体动态迁移、多实例并发、群组通信的特性。网络必须发展出能够理解并路由“身份标识”的新机制，实现身份与位置的解耦。

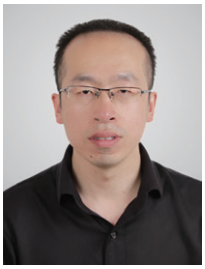
其次，网络需要具备语义感知与语义驱动的能力。未来的网络流量引导和转发不能只看IP和端口，还需要能识别流量的业务内涵和优先级，从而进行差异化的资源调度和服务质量保障，这是实现智能体高效协同的基石。

第三，内生安全必须成为网络的默认属性。智能体直接操作用户数据和关键业务，其通信链路必须是可信的。网络需深度融合零信任架构，实现以智能体身份为中心的动态微隔离、持续认证和加密传输，确保每一次交互都可验证、可审计。

再进一步看，Underlay网络本身需要完成“智能体化”的演进升级，作为提供网络差异化承载能力的“网络智能体”参与到业务工作流中，将差异化的SLA承载能力封装成不同的skills，与其他业务智能体通过MCP/A2A/RPCoS-LIM等协议进行交互，完成完整工作流的智能化闭环。

智能体互联网的构建是将智能从孤立节点解放出来，连接成智慧网络的过程。可以预见，未来的网络将不再是隐藏在应用之下的隐形基石，而是化身为智能体数字社会中有感知、会思考、能行动的“智能协作者”。智能体互联网的目标，不仅是一个更快的网络，更是一个能孕育和承载群体智能的新一代数字基础设施。ZTE中兴

网元内生智能架构及关键技术



武利明
中兴通讯数据产品资深系统工程师

段威
中兴通讯资深研发总工

全球通信网络正经历从“管道化”到“智能化”的范式变革。据IDC预测，全球AI算力已经正式迈入ZFLOPS时代，驱动网络流量呈现显著特征异变——智算中心间模型训练流量突发量可达日常流量的300%。AI推理请求的时延敏感度低于5ms，这种新型业务特征要求网络架构具备高实时性响应以及弹性调度的能力。

网络安全态势同样面临严峻挑战，DDoS攻击向高频、短时、分布方向发展，APT攻击的平均潜伏期已从2018年的107天缩短至现在的23天，迫切需要网元具备自主攻击检测与防御能力。

网络运维领域，Gartner研究指出，78%的网络故障修复仍依赖人工经验，平均故障定位时间超过4小时，这与5G-A网络要求的99.999%可用性形成尖锐矛盾，难以满足“零接触、零故障”的自治目标。

在此背景下，网络智能化正从边缘创新转向架构重构。3GPP R18标准首次将“内生智能”纳入网络切片管理框架，ETSI的ISG ZSM工作组已制定智能闭环控制的接口规范，这些进展标志着网络智能化已进入“算网智”深度融合的新阶段，需要将智能能力下沉至网元层级，才能突破传统架构的性能天花板。

网元内生智能架构

网元内生智能是实现网络智能化新范式的重要基础，需要具备弹性可扩展的智能化架构，包括统一的数据感知平台、分布式异构算力引擎、动态可分配的算力资源管理等，满足网络流量调优、安全运维、故障运维等不同领域的智能化演进需求。

中兴通讯路由器产品在网元传统架构上引入智能面，集成数据感知与智能化模型服务，并与网元业务协同实现“感知、规划、仿真、行动”智能化应用闭环流程，对不同的智能化应用提供公共基础设施，具备智能化应用快速开发与部署的能力。整体架构如图1所示。

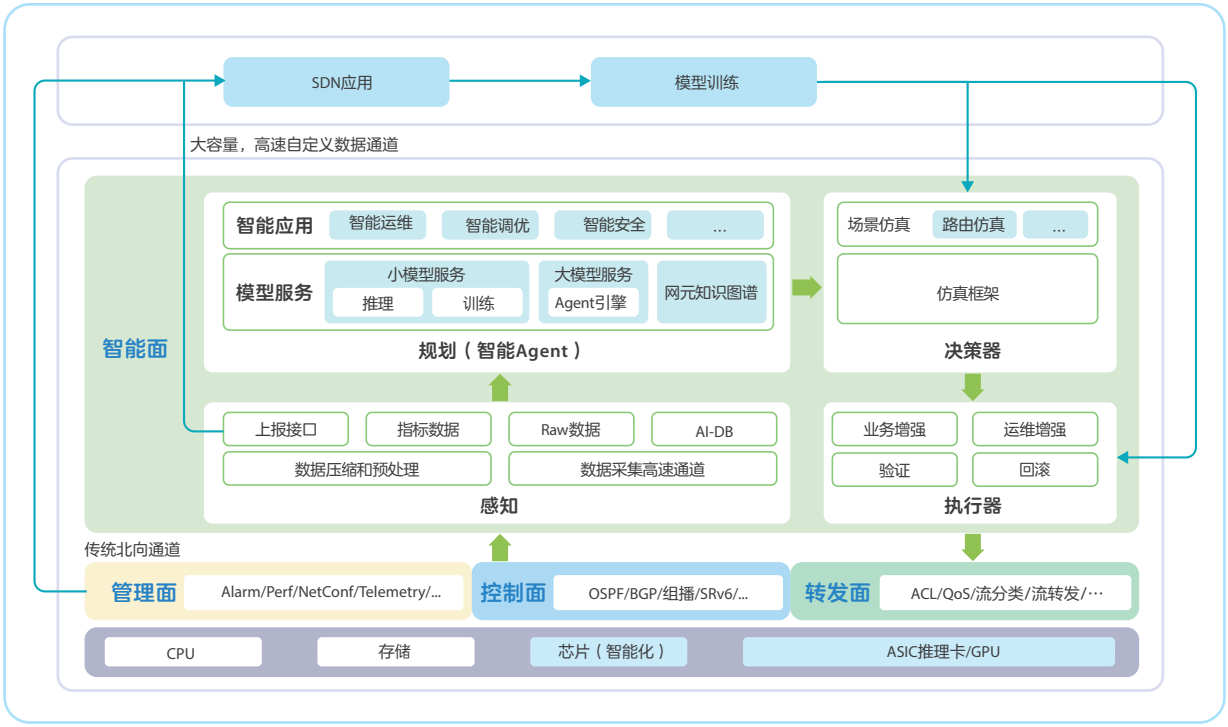
- 数据感知：构建统一的数据感知平台，实现KPI指标、告警信息及异构日志的统一采集与管理，在线卡、主控与AI算力单板之间设立独立高速通道，借助硬件机制实现毫秒级遥测数据采集与传输，通过数据订阅与推送服务，满足不同智能化应用对数据的差异化需求。
- 模型服务：智能面通过在主控与线卡集成AI芯片，实现高效本地计算，同时支持专用AI算力单板，满足高吞吐、高并发、高算力的智能化应用需求；支持轻量化AI模型的统一部署与算力分配，具备本地推理能力，并可开展在线增量训练与持续学习。

网元智能化应用

基于于网元内生智能化架构，通过AI赋能，网络可实现智能安全、智能调优、智能运维等网元智能化应用，提升网元的自主分析与决策能力。

智能安全

当前安全架构正向“纵深防御”与“零信任”方向演进，其中将安全能力下沉至网元层级成为关键环节。中兴通讯路由器产品基于网元内生智能架构实现了智能安全应用，基于AI驱动的异常检



▲图1 网元内生智能架构

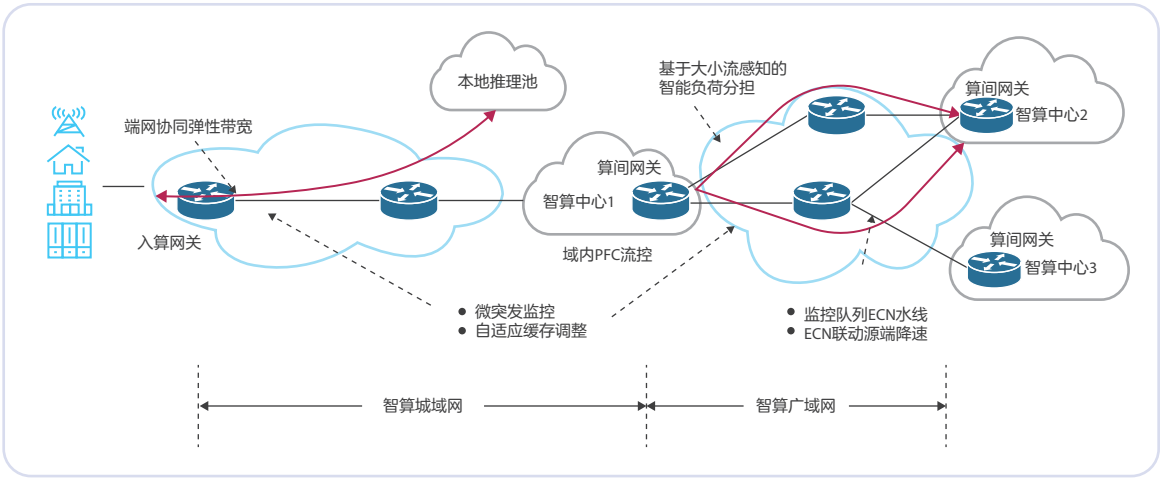
测与攻击模式分类模型，采用“应用会话行为异常检测+EDR（终端检测与响应）”协同机制，实现对APT与DDoS攻击的纵深防御，具备主动感知、自主学习、实时响应的全面安全防御能力。

- 应用会话行为异常检测系统：部署于主控板，对用户访问网元的会话行为进行实时监控，基于历史数据构建正常行为基线，识别异常登录、越权操作等可疑行为，并与EDR系统联动，实现APT攻击检测与防御。
- EDR系统：部署于专用AI算力单板，支持容器化灵活部署，持续采集操作系统层的进程、网络连接、日志等多维行为数据。通过AI/ML模型构建行为基线，精准识别异常执行路径、隐蔽信道通信、横向移动探测等攻击特征，实现对无签名攻击和LoL（Living-off-the-Land）类攻击的高检出率。

智能调优

当前跨数据中心智算业务具备典型的高突发、大带宽、低时延等特性，对广域网数据传输提出新的挑战，传统的QoS静态策略无法应对复杂多变的流量场景，需要网元设备具备智能化的流量识别以及自主流量调优能力。中兴通讯路由器产品在智能化架构基础上集成一系列智能调优技术，相互协同，实现智算场景的无损转发（见图2）。

- 端网协同弹性带宽：通过端侧与入算/算间网关协同，基于业务带宽需求（如SLA合约请求），动态建立满足业务需求的弹性带宽转发路径，实现带宽资源按需分配。
- 自适应缓存调整：基于微突发流量监控，实时动态调整本地缓存策略，吸收突发流量，避免丢包，本地局部优化实现无损，减少对PF流控的依赖。
- ECN联动源端降速：网元内嵌智能QoS技术，对应用流量进行建模分析，结合时序预测算法（如ARIMA、LSTM）预测网络流量趋势，并利用强化学习等方法自适应优化



▲ 图2 网络智能调优架构

ECN阈值；当网络出现拥塞时，由网元标记ECN并通知源端主动降速，实现拥塞前导式调控。

- 域内PFC流控：在ECN机制未能及时缓解突发流量的情况下，启动域内PFC实现流量控制机制，防止拥塞扩散，保障关键业务不丢包。
- 智能负荷分担：基于流量生命周期、大小分布等动态特征，优化多路径负载均衡策略，实现更精细化、更均衡的流量调度。

智能运维

网元自动化故障诊断能力是实现网络分钟级故障自愈闭环的核心。中兴通讯路由器产品基于网元智能架构，在智能面融合时间序列异常检测、故障知识图谱、故障分类模型与故障诊断思维链等多种故障诊断算法模型，实现故障的精准感知、快速定界与智能处置。

- 时间序列异常检测：实时监控网元KPI、流量等时序数据，基于无监督算法模型建立KPI与流量的基线模型，实现对异常突变的快速检测与预测。
- 网元故障知识图谱：基于业务逻辑与故障传播关系构建网元内部故障依赖图谱，结合故障分类模型，提升故障定界定位的准确性与可解释性。
- 故障分类模型：利用历史标注的故障数据，通过机器学习与深度学习方法训练分类模型，实现对典型故障类型的自动识别。
- 故障诊断思维链：将专家经验与标准化排障流程结构化，构建可执行的诊断推理链，实现从告警到根因的自动化定位闭环。

网元内生智能作为高性能智算网络演进的核心范式，通过将轻量化AI能力深度集成于网络单元，实现了网络架构从被动响应向自主感知、自适应优化与自治运维的跨越式升级。本研究基于中兴通讯在智能安全、流量动态调优及智能运维领域的技术实践，初步构建了“感知–决策–执行”闭环体系：在智能安全方面，通过应用会话行为分析与EDR协同机制，实现对APT攻击的毫秒级检测与闭环处置；在动态调优领域，创新性地融合端网协同弹性带宽、自适应缓存调整与ECN联动源端降速等技术，有效保障高动态业务场景下的确定性传输；在智能运维层面，通过时间序列异常检测、故障知识图谱与思维链诊断的多技术融合，将故障定位效率提升至亚秒级。实验结果表明，该架构可使网络控制时延降低60%以上，故障自愈覆盖率突破85%，为构建高可靠、自适应的通信基础设施提供可落地的技术路径。ZTE中兴

数据快递与AI入算业务使能技术——高性能广域网（HP-WAN）

黄光平
中兴通讯有线标准总监

熊泉
中兴通讯分组网络标准预研工程师

随着国家“东数西算”战略的实施部署，以及生成式人工智能（GenAI）与高性能计算（HPC）的高速发展，算力中心承载的数据量与协同需求呈指数级增长，跨地域算力资源的实时调度与海量数据传输已成为关键挑战。此外，在AI/HPC跨地域协同的多种场景中，包括训练前模型与数据在数据中心间的快速上载、训练期间跨设备数据状态同步、科学数据快递及灾备传输等，这些任务不仅对传输速率和时延有极高要求，也对数据完整性和系统长期稳定运行提出了更高性能要求。由于传统广域网在长距离传输中面临带宽利用率低、时延不可控等瓶颈，导致算力协同效率受限，急需一种以“高带宽、低时延、无损化”为特征，在高带宽利用率下提供有效高吞吐的广域网技术，其既能实现跨域算力资源的毫秒级联动，又能保障海量数据在长距离下的传输效率，使存算分离、多中心分布式训练等场景突破地理限制。

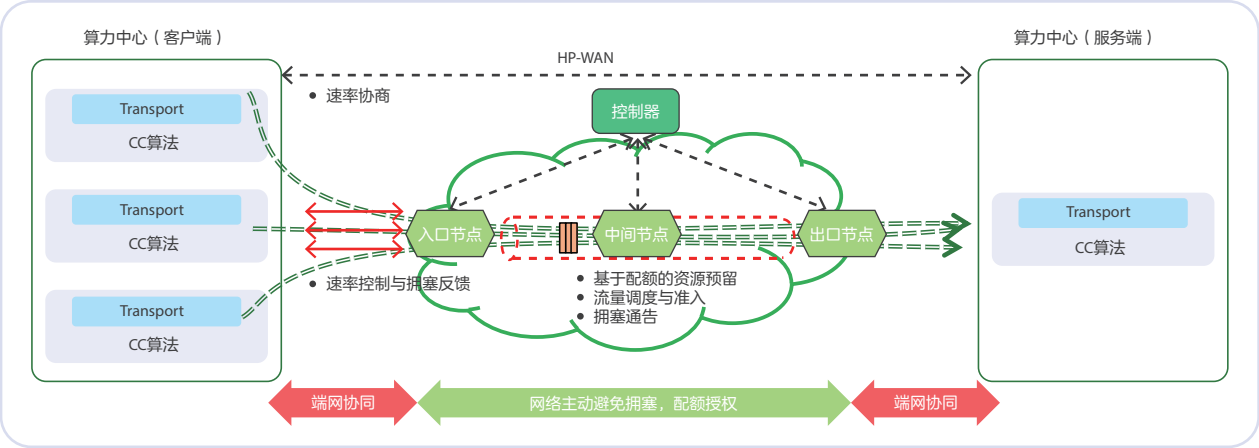
高性能广域网概念

高性能广域网（high-performance wide area network，HP-WAN），以跨站点或跨数据中心构建的广域网为基础，满足AI/HPC对高速率、低延迟和高可靠性的苛刻传输需求，为高速、低延迟和超高容量应用而设计的广域网高通量技术，聚焦基于网侧增强的端网协同方案。网络进行主动拥塞避免，在端网之间进行流量及资源调度，通过双向交互协商速率等保障高性能数据传输，在保障资源利用率和公平性的同时，实现高通量传输。

HP-WAN作为相较于传统WAN的技术演进方向，承担起面向多种算力互联场景下的关键数据高效承载的职责：面向秒级至分钟级的任务完成时间目标，致力于提供超高有效吞吐量（即大容量数据在限定时间内完成传输的能力），并在提升带宽利用率的同时，确保链路资源在多业务并发下的公平共享与服务质量保障，避免多流竞争导致的慢流拖尾及FCT（flow completion time）传输抖动。

高性能广域网架构及关键技术

如图1所示，为了保障大容量数据的广域网高性能传输，HP-WAN在IETF（参考draft-xhy-hpwan-framework）提出基于端网协同的架构，提供任务式应用需求与端侧主机协商双向速率，通过向端侧主机动态分配和授权发送流量的配额来防止拥塞，同时实现基于配额的资源调度、准入控制和流量控制，满足最优完成时间目标。通过端网协同速率协商，客户端和服务端能够以更精细的方式高效快速地调整发送速率，避免端侧传输协议被动调速，提高广域传输吞吐量；通过基于配额的资源预留，网络增强对流量的调节能力和资源调度能力，保障所有任务的传输需求及资源保障；通过流量调度与准入，实现多流之间网络带宽资源的合理分配及动态调度，控制智算



▲图1 HP-WAN端网协同架构

业务传输的最大及最小速率，在满足高吞吐要求的同时，避免慢流拖尾现象，实现集合通信多流传输的同步性。同时，采用快速拥塞感知及通告机制，在网络拥塞发生时，网络可以在近源端进行快速拥塞反馈，能够迅速且准确地对流量速率进行反馈通告，并缓解网络拥塞。

根据端侧主机与网络的协同，HP-WAN对端侧主机/入口节点/中间节点/控制器等提出以下相关功能要求：

- 端网协同速率协商：端侧与网络进行流量规划及调度，根据流量传输需求协商速率；
- 网络动态资源预留：网络需要提供任务感知及资源调度，保障所有任务的传输需求及资源保障；
- 流量调度及准入：网络对端侧流量请求进行授权准入；
- 网关拥塞通告：网络拥塞节点可向代理网关节点发送快速拥塞通告报文，再由代理网关节点向发送端设备发送拥塞通告报文。

端网速率协商

智算业务中突发的大容量数据流量传输可能导致网络内的瞬时拥塞、丢包和排队延迟。在拥塞控制机制中，端侧对网络的带宽资源无感知，导致调速不平滑，吞吐量下降。因此，高性能广域网将向端侧协商速率策略，从而实现速率协同及拥塞避免。

根据端侧主机与网络的协同机制，HP-WAN在IETF（参考draft-xiong-hpwan-signaling-solution）提出3种速率协商策略：

- 最优速率或最优速率传输：网络为大容量数据提供资源调度机制获得QoS保障，实现最优速率传输，端侧主机可按照协商的最优速率或最优速率范围传输。
- 最小速率传输：网络为大容量数据提供最小的资源预留保障，实现最小速率传输，端侧主机可按照不小于协商的速率传输。
- 最大速率传输：网络为大容量数据提供资源预留的上限，实现最大速率传输，端侧主机可按照不大于协商的速率传输。

基于配额的动态资源预留

在HP-WAN场景中，数据传输有任务式传输的需求，且任务有预期性，需要提供任务感知及资源调度，保障所有任务的传输需求及资源保障。HP-WAN在IETF（参考draft-xiong-teas-rsvp-resource-quota）提出分布式信令的方式实现基于配额的动态资源预留机制。基于配额（quota）的调度是一种资源管理策略，配额可定义为一定时间内的可用资源（带宽、队列、buffer等），网络可根据任务需求分配和授权配额，并且实现基于配额的资源调度，进行主动拥塞避免，保障基于

配额及其速率的高效转发。同时，HP-WAN对于配额资源需要基于速率控制进行动态调度，通过实现多流之间网络带宽资源的合理分配及动态调度，控制智算业务传输的最大及最小速率，协同端侧传输协议进行业务流量调度，在满足高吞吐要求的同时，避免慢流拖尾现象，实现集合通信多流传输的同步性。

流量调度及准入

HP-WAN在IETF（参考draft-xhy-hpwan-framework）提出可在接收流量后基于协商速率进行流量调度与策略执行，包括对流量分类、按业务类型区分优先级、提升关键流量QoS等级、对流量进行整形，例如聚合小鼠流（mouse flows）或分片大象流（elephant flow）等。网络入口的流量调度策略执行可规范数据流，而流量根据网络可用资源进行准入控制可以消除拥塞并最小化流完成时间。为了支持端网速率协同，网络可扩展RSVP-TE协议进行基于速率控制的动态资源调度和准入，通过预留最小速率对应的最小带宽配额保障单流完成时间，通过动态调度最大速率对应的最大带宽避免多流竞争导致拥塞丢包。网络节点应基于协商的QoS（服务质量）与速率执行准入及流量控制。通过准入控制与拥塞控制的结合，可在高效利用网络容量的同时，实现高吞吐量与低完成时延。

网关快速拥塞通告

HP-WAN在IETF（参考draft-xiao-rtgwg-proxy-congestion-notification）提出拥塞节点可向代理网关节点发送快速拥塞通告报文，再由代理网关节点向发送端设备发送拥塞通告报文。HP-WAN需要为每一个端侧设备指定一个用于快速拥塞通告的代理网关节点，代理网关节点应知晓端侧设备所能够解析的拥塞通告报文。代理网络节点通过IGP协议或BGP协议向外通告自身的拥塞通告代理能力及所代理的端侧设备的IP前缀，网络中的设备收到代理网络节点的通告后，记录代理网络节点与其所代理端侧设备的映射表，网络中一旦发生拥塞，检测到拥塞的网络节点通过拥塞报文的源IP地址找到代理网络节点，可扩展ICMP或UDP协议向代理网络节点发送快速拥塞通告报文，再由代理网络节点向发送端设备发送拥塞通告报文。

总结与展望

标准推进方面，IETF标准组织针对低时延、高吞吐、低丢包等智算场景需求，已在传输、管控、路由等领域进行相关标准讨论。例如传输领域的SCONE、TSVWG、CCWG等工作组针对RDMA包括RoCE等与TCP、QUIC等协议的适配，CUBIC及BBR等拥塞控制算法的优化等进行了讨论。针对AI/HPC等广域网大容量传输需求，IETF WIT域已于2024年7月开始讨论高性能广域网场景及需求等，并于2024年11月成功召开HP-WAN BOF，明确了广域网需要满足高速、低延迟与超高容量的应用场景，及高吞吐低时延等基础需求，HP-WAN架构及其关键技术相关标准已成为面向智算场景的研究热点及标准化方向。

技术趋势方面，高性能广域网中无损技术与有损技术并存，广域无损技术能够为业务提供低延迟、低丢包和高带宽利用率的数据传输服务，除光互联方案之外，确定性网络技术也可用于提供广域长距无损承载能力。由于广域无损对网络有极高要求，对于时延不敏感的业务，也可增强网络能力提供广域容损的数据传输服务。基于IP的高性能广域传输方案能够以更低的成本支持更长的传输距离，基于网侧主动拥塞控制和配额协商，进一步增强端网速率协同，具备满足大容量限时传输的广域高性能传输需求的潜力。

在技术迭代和市场需求的双重推动下，高性能广域网将逐步替代高成本网络专线，成为支撑未来多场景低时延、高可靠、高安全网络连接的主力军。ZTE中兴

Scale-Up互联技术

潘文斌
中兴通讯数据中心网络架构师

随着大语言模型（LLM）参数规模从千亿级向万亿级甚至十万亿级的爆发式演进，传统单机8卡XPU服务器的计算资源与显存容量承载瓶颈日益凸显，必须使用大规模服务器集群进行训练。随着集群规模增大，单纯扩大数据并行（DP）维度面临上限。为了继续扩大集群，需要引入张量并行（TP）和流水并行（PP）。当并行域（如TP>8）超出单台服务器的范围时，跨服务器张量并行（TP）成为必然选择，而跨设备的TP All-Reduce通信成为制约大规模分布式训练性能提升的主要瓶颈。同时随着混合专家模型（MoE）在Transformer架构LLM中的规模化应用，更使跨服务器专家并行（EP）成为分布式训练和推理的关键技术需求，而跨服务器的All-to-All通信成为新的瓶颈。

为应对TP和EP对网络带宽与延迟极为严苛的要求，纵向扩张网络（Scale-Up）成为业界主流技术路径。通过Scale-Up网络，可将几十、上百甚至上千张XPU高速互联，构建为超节点（SuperPoD），像一台超级XPU服务器一样实现高效的计算和通信协同能力。Scale-Up不是简单地将多卡进行硬件堆砌，而是需要超高带宽、低延迟的互联技术进行构建。

不同于横向扩张网络（Scale-Out）已经基于Infiniband和RoCEv2形成了业界共识，Scale-Up当前还没有一个统一的标准，呈现出一超多强的局面。NVLink在英伟达（NVIDIA）的Scale-Up垂直整合方案中广泛应用，但依赖单一供应商专有技术所带来的高昂成本和封闭生态、深度的厂商锁定、有限的供应选择，已成为AI基础设施发展的沉重负担。在此背景下，国内外都涌现出多个Scale-Up技术方案：AMD将迭代多年的Infinity Fabric技术开放共享，促成UALink联盟的成立；Broadcom推出面向Scale-Up场景的SUE（Scale-Up Ethernet）以太网方案，并已被ESUN（Ethernet for Scale-Up Networking）确立为该工作组的推荐传输协议；国内字节跳动、北京大学联合设计Ethlink，阿里巴巴、中国科学院计算技术研究所牵头组成ETH+联盟，腾讯、中国信通院在ODCC立项ETH-X，还有设备制造商和芯片提供商推出各自的专有Scale-Up互联协议。

接下来我们从AI网络视角介绍这些“后NVLink时代”的Scale-Up主流技术流派。

UALink

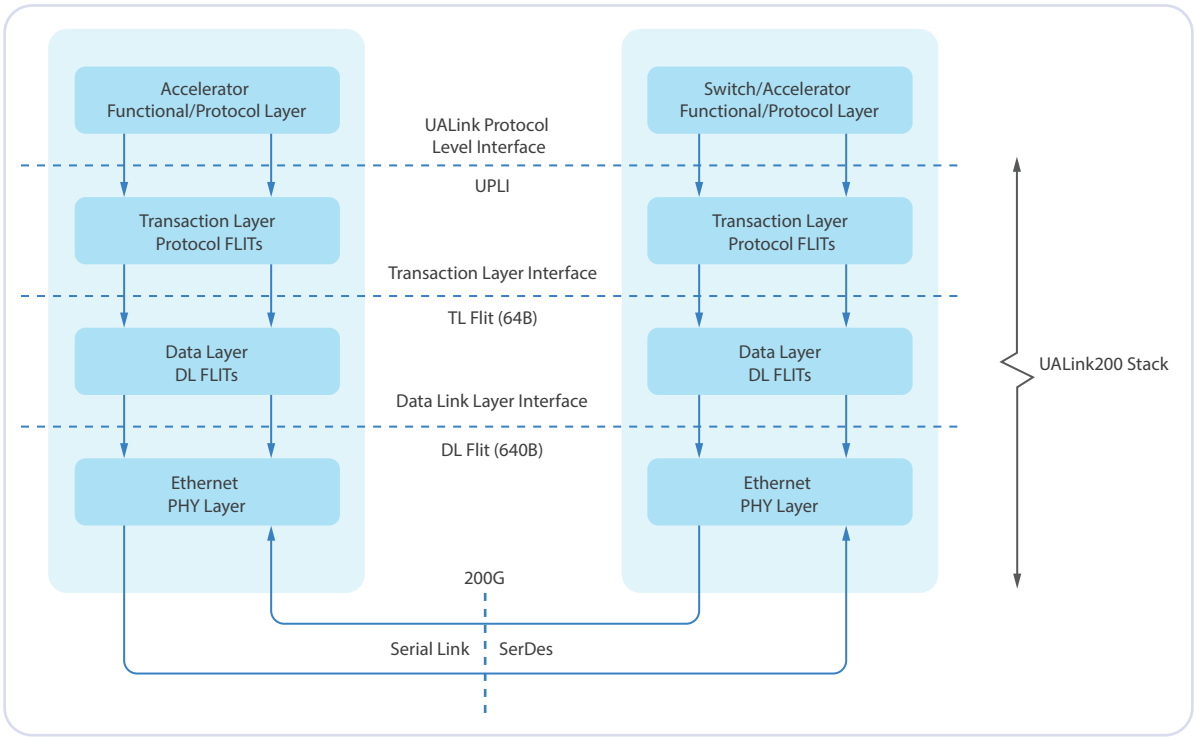
UALink协议由UALink（Ultra Accelerator Link）联盟推动。UALink联盟由AMD牵头，联合行业头部硬件厂商、芯片设计企业，共同参与标准制定与生态共建。

UALink目前有两种规范：基于Ethernet物理层的200G规范，以及基于PCIE物理层的128G规范。

200G规范定义四层协议栈：物理层（PL）、数据链路层（DL）、事务层（TL）、协议层（UPLI）（见图1）。

物理层基于标准的802.3以太网物理层，单个UALink通道的最大数据传输速率为200Gbps，也可以降速为100Gbps使用，链路的最大通道宽度为4通道，单个UALink station的最大传输带宽为800Gbps。

数据链路层位于事务层与物理层之间，将事



▲ 图1 UALINK 200G协议栈

务层下发的64B Flits封装为适配物理层的640字节Flits。该层支持链路级重传功能（link layer retry），以640B Flits为基本单位实现，若接收端CRC校验失败，会向发送端DL发起重传请求。

事务层负责将两个UPLI接口（Originator/-Completer）接收方向通道的协议传输转换为以64B为单位的UPLI Flit。此外，事务层还会将从数据链路层接收到的TL Flit，重新转换为UPLI接口上的UPLI通道事务。

协议层定义了一套逻辑信令接口与通信协议，设备可基于此，通过各类请求报文和响应报文完成数据与控制信息的交互。UPLI协议具备原生的灵活扩展能力，支持厂商为同类型加速器间的通信定制私有协议报文。UPLI支持单系统内最多1024个加速器或端点，通过10位标识符完成互联通信。UALink交换机依托这10位加速器源标识符与目的标识符，在发送端和接收端之间转发请求与响应报文。

UALink 200G基于以太网物理层实现，更偏向于GPU之间互联，不支持异构计算设备混联。UALink 128G可以通过PCIE支持混接GPU、CPU、存储。

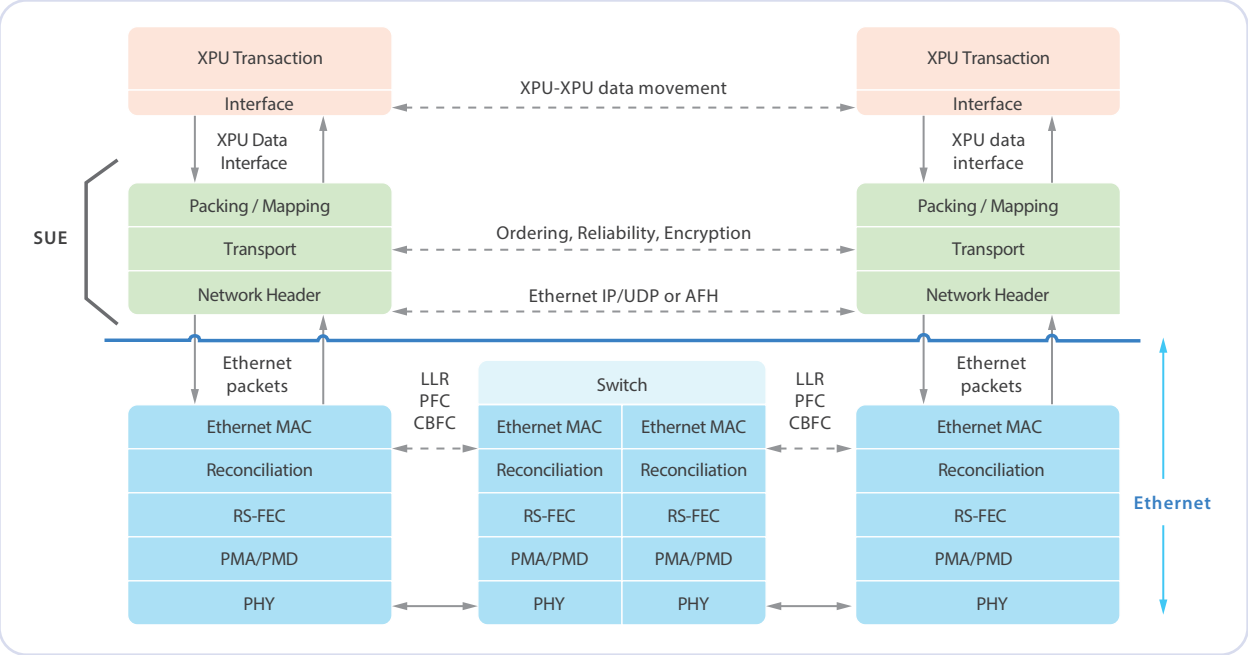
200G和128G的1.0规范已正式发布，多厂商联盟持续推进生态落地，已形成芯片、IP核等配套解决方案，逐步缩小与NVLink的性能差距。

SUE

SUE协议由博通（Broadcom）主导设计，依托博通在交换芯片领域的技术积累与行业资源推进。

SUE设计目标是支持1024个XPU，采用类AXI的双工数据接口，通过虚拟通道（VC）将事务映射到不同流量类别。其协议栈分为三层（见图2）：

- 映射打包层：将发往同一目标（destination，VC）的事务聚合成最大4096B的SUE协议字节单元。
- 传输层：添加可靠性头部（RH），包含序列号（PSN）、虚拟通道（VC）以及确认机



▲ 图2 SUE协议栈

制（RPSN），并添加CRC检验。

- 网络层：支持多种报文头，包括标准以太网IP/UDP报文头、优化的AI转发报文头（AFH Gen1）以及高度压缩的AFH Gen2（6B~12B）。

SUE提供三类接口：命令接口，支持FIFO信用机制、AXI4总线，用于传输事务指令和数据（包含操作码、长度和目标XPUID）；管理接口，基于AXI的寄存器配置通道；以太网接口，支持200G/100G速率。

ESUN

2025年OCP全球峰会（Open Compute Project Global Summit）期间，由AMD、英伟达、博通、Meta等12家国际厂商联合发起成立ESUN（Ethernet for Scale-Up Networking）工作组，依托OCP组织进行开放治理。ESUN工作组处于规范制定与整合阶段，目标是构建覆盖全数据中心的开放以太网Scale-Up协议栈。参与工作组的成员包括北美AI行业芯片/硬件厂商、大云厂商、开源/闭源大模型厂商等所有重要玩家，未来有望形成产业联盟。

ESUN主要聚焦于网络层和数据链路层，上面的传输层则由SUE-T或其他协议负责。

ESUN定义了新的报文头，将20B的IP头去掉，定义了4B的EH Header，其中包括EH-ECN（和传统IP字段的ECN相同）、EH-QOS（精简版的DSCP）、可用于负载均衡的Flow Label。

EthLink

EthLink（Ethernet Link）由字节跳动牵头设计研发。《字节跳动GPU Scale-Up互联技术白皮书》已经发布，EthLink已在字节跳动AI场景中完成实践验证。EthLink最大支持单跳1024个XPU互联。

EthLink首先对协议栈进行了改造，支持RDMA和L/S双语义：TMA通过DMA Read和DMA Write语义完成Global Memory和Shared Memory之间的数据传输，LSU通过Load/Store语义完成Shared Memory到寄存器之间的数据传输。

针对传统IP报文臃肿的问题，EthLink设计了极致优化的报文头——OEFH（Optimized EthLink Forwarding Header）。同时构建了一套专为GPU间通信设计的、更轻量级的链路层和事务层协议，使用6B的OEFH进行寻址和转发，进而大幅提升GPU间通信的有效Payload率。EthLink封装和OEFH报文头如图3所示。

ETH+

ETH+由中国科学院计算技术研究所与阿里巴巴牵头，联合行业伙伴成立ETH+（高通量以太网）联盟推进。ETH+的《Scale-Up互连协议白皮书》已经发布，正在推进标准细化与技术落地。

ETH+引入语义适配层，用于桥接上层加速器操作语义与基础网络层之间的差异，将高层的通信操作（如Load/Store、RDMA Read/Write、Send/Receive等）映射为统一的基础通信语义。为高效实现集合操作，协议提供了Reduce、Broadcast、ReduceScatter、AllGather、AllToAll等集合操作专用语义，并将其卸载至Scale-Up网络的交换机或网络接口执行，减少开销。

ETH+基础网络层在协议包头、链路层功能、FCS方面遵循“极简设计”理念：

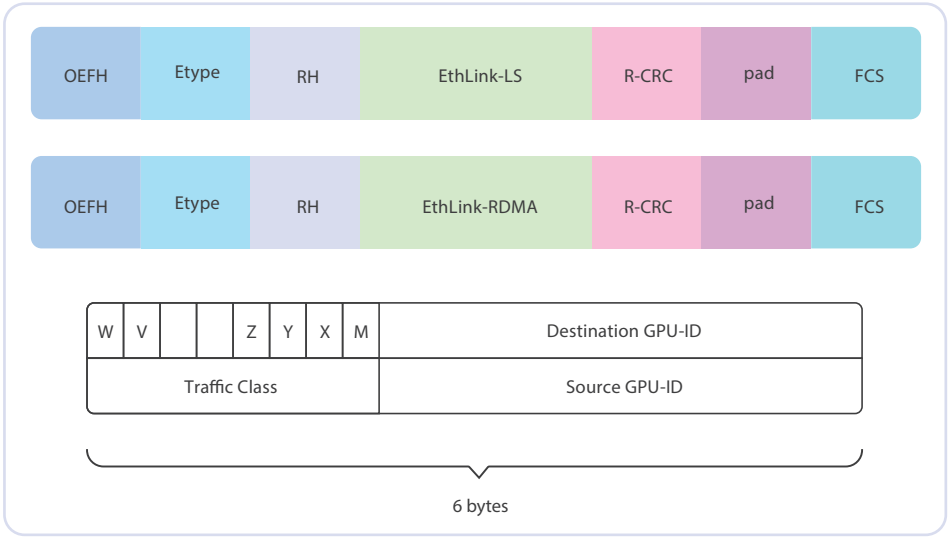
- 前导码压缩到1B，填充为1010xx11，分别用于同步发送/接收的时钟、映射不同帧类型、标识帧数据起始位置；
- 链路层功能旁路：取消了D_MAC、S_MAC、Type/Len字段，取而代之的是UID base header字段；链路层设备之间的数据帧转发通过识别UID编号而非MAC地址实现；
- FCS简化：ETH+协议允许完全移除帧校验序列（FCS）机制。

ETH-X

ETH-X由中国信通院与腾讯牵头，在ODCC（开放数据中心委员会）立项推进，联合行业伙伴共同研发。目前已完成核心架构与协议规范定义，1.0版本规范已经在ODCC发布，依托ODCC平台推进行业适配。

ETH-X分层架构如下：

- Scale-Up访存协议定义GPU芯片−GPU芯片、GPU芯片−内存池模组的事务访问方式；
- Scale-Up互通协议定义GPU芯片−Switch芯片间的数据包高效可靠传输；
- Scale-Up D2D互联定义计算DIE与IO DIE的物



▲ 图3 EthLink封装和OEFH报文头

未来，在NVLink主导的NVIDIA超节点场景之外，Scale-Up领域的发展走向充满变数：或许国内统一为CLink通用技术标准，或许网络层将由ESUN实现归一化整合，亦或是技术与利益分歧持续存在，长期维持多元竞逐的格局。

理互联方式。

报文头优化方面，ETH-X设计了新的PRI（packet rate improvement）统一转发头，替代传统以太网的DMAC和SMAC域，共12字节，其中前2个字节为Network DeviceID域，作为网络地址用于网络路由字段，其余10字节为设备地址（User Defined Address），用于设备内部寻址，网络转发节点忽略其内容。

CLink

在国家工信部指导下，中国电子标准化研究院会同北京市经信局联合政产学研用各方，共同发起了计算互联总线协议（CLink）联合倡议，核心目标是提升计算产业全链协同能力。CLink以开源、共研、共享为模式，打造统一的计算互联总线协议标准簇，涵盖总体架构、通信语义、流量控制等关键模块，为大规模智算集群提供统一技术规范。其价值主张为降低集群互联成本与适配周期，提升规模化部署效率，凝聚全产业链共识。

目前CLink已形成初步标准体系，以开源开放智算互联协议为蓝本持续迭代，获得国内产业链广泛关注，正在推进实际场景落地应用，致力于构建自主可控的计算互联生态。

总结

为破解NVLink“一超独大”的格局，业界涌现出了多种Scale-Up方案。主流方案不约而同地选择了基于以太网演进的技术路线，且均针对以太网封装报文头开展了针对性优化，形成了面向AI算力集群的开放互联技术共识。然而，共识之下，开放阵营内部的利益博弈与技术分歧尚未消弭，整体呈现出百花齐放、多元共生的发展态势。

UALink依托多厂商联盟的协同力量，稳步推进互联标准的落地与推广；SUE凭借博通在交换芯片领域的深厚技术积淀，构建起专属的硬件生态优势；ESUN借助OCP的行业组织号召力，统筹跨厂商完成规范的制定与整合；而EthLink、ETH+、ETH-X等方案，则依托国内头部企业与学术机构的实践积累和技术创新能力，走出了独具特色的技术发展路径；CLink由政府部门指导，凝聚全产业链广泛的关注和共识。

未来，在NVLink主导的NVIDIA超节点场景之外，Scale-Up领域的发展走向充满变数：或许国内统一为CLink通用技术标准，或许网络层将由ESUN实现归一化整合，亦或是技术与利益分歧持续存在，长期维持多元竞逐的格局。无论最终走向如何，对于网络领域的科研与工程从业者而言，都无疑是一场值得期待的技术盛宴。ZTE中兴

基于GSE技术的十万卡级组网：智算中心Scale-Out网络新路径

王恒
中兴通讯交换机产品规划经理

在人工智能技术浪潮的驱动下，智算数据中心迎来跨越式发展机遇。大模型训练与推理对算力的需求呈指数级激增，推动智算中心从小规模集群加速向超大规模集群演进，十万卡级GPU组网已成为行业竞争的核心基础设施壁垒。但海量GPU节点的高频数据交互使网络带宽需求同步激增，通信效率成为制约训练效率的关键瓶颈。

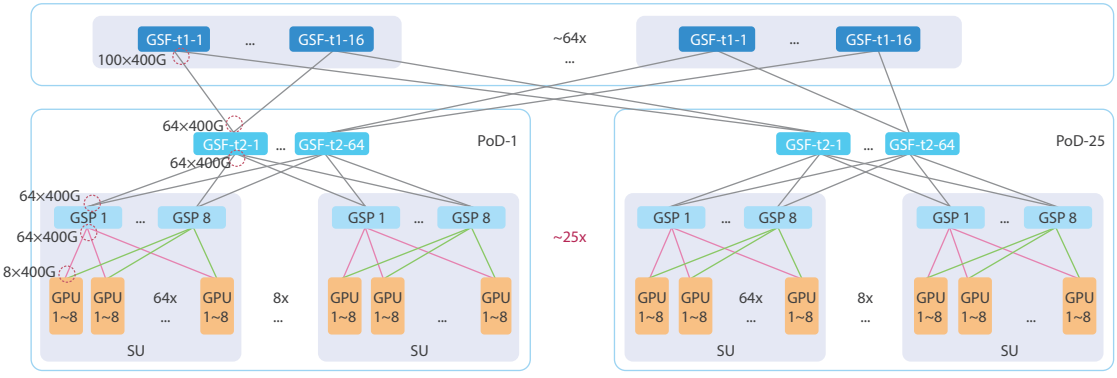
为破解当前网络架构中算力与网络的适配难题，中国移动联合产业伙伴自主创新研发出全调度以太网（Global Scheduling Ethernet，GSE）技术体系。该技术通过深度优化以太网架构，构建“主动调度+精准分发”的传输机制，为智算数据中心Scale-Out网络提供全新研究视角。

相比RoCE技术，GSE技术创新提出报文容器（packet container，PKTC）、全局动态调度队列（dynamic global scheduling queue，DGSQ）等概念。通过对报文容器的转发与逐容器喷洒，实现单流在多路径上的均匀分担，大幅提升带宽利用率；通过DGSQ搭建拥塞控制体系，引入授权请求与全局调度机制，确保流量负载不超过网络承载上限，从根源上规避拥塞丢包。采用GSE技术搭建十万卡级网络成为当下热点研究对象。GSE组网架构依据互联方式的差异可划分为三层组网架构和多PoD互联架构两类，采用层次化、模块化方式设计，适配不同应用场景需求。

三层组网架构

GSE三层组网架构采用搭载51.2T芯片的盒式交换机搭建，单台交换机可提供128个400GE端口，按无收敛方式设计。

根据层次化思想，网络分为三层：接入层、转发层、互联层。如图1所示，接入层由GSP



▲图1 GSE三层组网示意图

（global scheduling processor，全调度以太网处理节点）设备构成，用于将GPU服务器接入网络；转发层由GSF-T2（global scheduling fabric，全调度交换网络）设备构成，用于PoD（point of delivery）内GSP之间的转发；互联层由GSF-T1设备构成，用于PoD间的互通。

根据模块化设计，网络分为SU（scale unit，扩展单元）、PoD、DC（data center，数据中心）等不同层级的模块。8轨部署中的8台GSP设备及其所连接的服务器共同构成一个SU模块；多组SU及其共同相连的GSF-T2设备组成一个PoD单元；多个PoD及其之间互联的GSF-T1设备组成一个DC网络。

组网部署

每台服务器搭配8张GPU卡，提供8个400GE端口，按需选择零轨部署或轨道化部署模式。GSP设备独立部署，其中64个400GE端口用于与服务器互联，剩余64个端口与64台GSF-T2设备互联。GSF-T2设备独立部署，其中64个400GE端口通过Full Mesh方式与GSP设备互联，另外64个端口与多平面部署的GSF-T1设备互联。GSP-T1设备独立部署，为提高组网规模，采用多平面设计架构。最大可分为64个平面，每个平面内部署16台GSF-T1设备，分别与其他PoD中对应序号的GSF-T2设备互联。单个PoD可提供4K卡规模部署，三层组网架构最大可接入128PoD，整网GPU卡容量最大可达512K。若需支撑10万卡规模部署，仅需配置25个PoD单元即可满足需求。

功能部署

GSP设备、GSF-T2设备、GSF-T1设备之间部署EBGP协议，通过BGP扩展属性将GSE所需的DPORT信息、GSPID信息同步发布至整个网络。

部署GSE N2N功能，GSP设备作为进入/退出GSE域的接入点，部署基于容器的转发策略、GSE头信息的封装与解封装、授权请求与响应处理及数据排序等功能；GSF设备作为GSE域内转发节点，部署GSE流量识别、基于GSE头信息的多路径转发等功能。

GSP设备与服务器相连的端口按需部署PFC（priority-based flow control，基于优先级的流量控制）功能，当源端口+优先级对应的缓存使用量到达阈值时，可向源服务器发起PFC反压，实现源端速率调控。

流量模型

同GSP设备下转发，源服务器网卡发送IP报文，经GSP设备，直接以IP报文形式转发至同一GSP设备下的目的服务器网卡。

同PoD内服务器间转发，源服务器网卡发送IP报文，经源GSP设备封装GSE头信息，采用容器喷洒方式至GSF-T2设备，GSF-T2设备根据GSE头信息转发至目的GSP设备，目的GSP设备解封还原为IP报文，转发至目的服务器网卡。

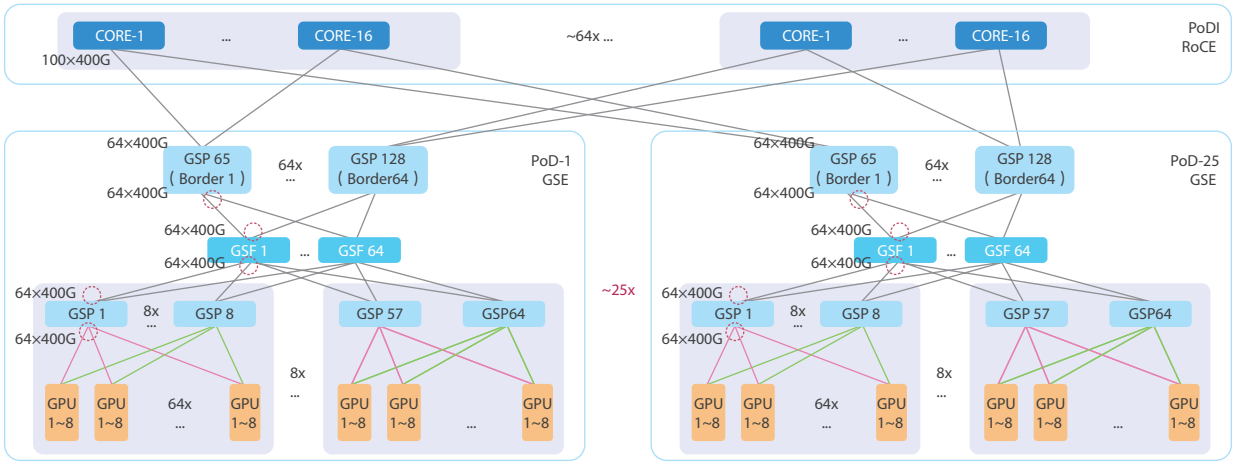
跨PoD转发，源服务器网卡发送IP报文，经源GSP设备封装GSE头信息，容器喷洒至GSF-T2设备，GSF-T2设备根据GSE头信息转发至GSP-T1设备，GSF-T1设备进一步转发至目标PoD的GSP-T2设备，最终经目标PoD的目的GSP设备解封还原成IP报文，转发至目的服务器网卡。

多PoD互联架构

GSE多PoD互联架构同样采用搭载51.2T芯片的盒式交换机搭建，单台交换机提供128个400GE端口，按无收敛方式、层次化、模块化设计。

组网部署

如图2所示，每台服务器搭配8张GPU卡，提供8个400GE端口。按需选择零轨部署或轨道化部署模式。GSP设备独立部署，其中64个400GE端口与服务器互联，剩余64个端口连接至64台GSF设备。GSF设备独立部署，其中64个400GE端口通过Full Mesh模式与GSP设备互联，另外64个端口与BORDER设备互联。BORDER设备独立部署，



▲ 图2 GSE多PoD互联（Border方式）组网示意图

其中64个400GE端口采用Full Mesh方式与GSF设备互联，剩余64个端口连接至用于PoD间互联的CORE设备。CORE设备独立部署，为扩大PoD间互联规模，采用多平面方式部署。共设置64个平面，每个平面内部署16台CORE设备，分别与其他PoD中对应序号的BORDER设备互联。单个PoD可提供4K卡（400G带宽/GPU）规模部署，最大可实现128个PoD互联，整网GPU卡容量最大可达512K。支撑10万卡规模部署时，需配置25个PoD单元。

功能部署

PoD内采用GSE N2N方式部署，PoD间采用IP RoCE方式部署。

对于PoD内，GSP设备、GSF设备、BORDER设备之间部署EBGP协议，通过BGP扩展属性发布GSE所需的DPORT信息、GSPID信息。部署GSE N2N功能，GSP设备和BORDER设备作为接入/退出GSE域的接入点，GSF设备作为GSE域内转发节点。GSP设备与服务器之间按需部署PFC功能，当拥塞时可以向源服务器网卡发起PFC反压，实现源端速率调控。

对于PoD间，BORDER设备与CORE设备之间部署动态路由协议（如EBGP协议），打通网络三层路由可达性。部署RoCE功能，配套部署PFC、ECN功能，为无损通信提供保障。

流量模型

同GSP设备下转发，源服务器网卡发送IP报文，经GSP设备，直接以IP报文转发给同一GSP下的目的服务器网卡。

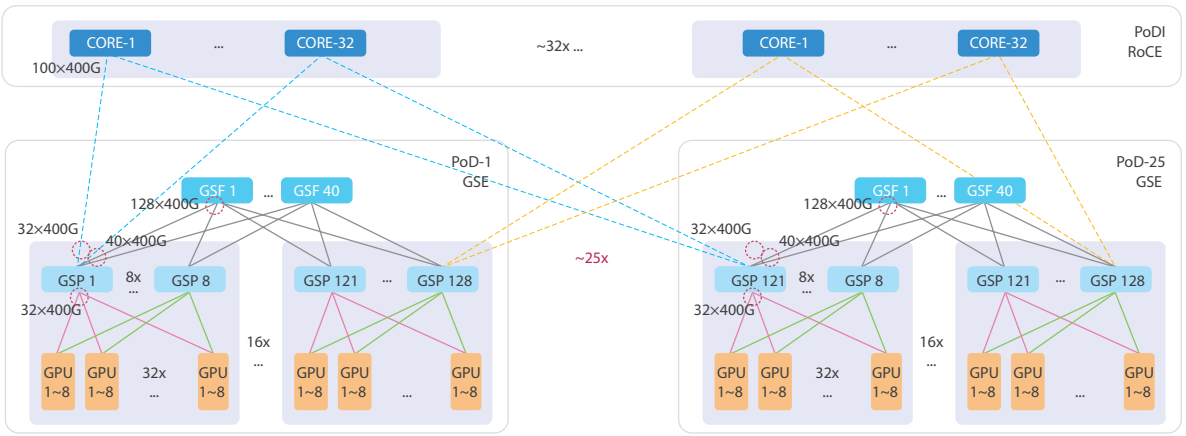
同PoD内服务器间转发，源服务器网卡发送IP报文，经源GSP设备封装GSE头信息，容器喷洒至GSF设备，GSF设备根据GSE头信息转发至目的GSP设备，目的GSP设备解封还原成IP报文，转发至目的服务器网卡。

跨PoD转发，源服务器网卡发送IP报文，经源GSP设备封装GSE头信息，容器喷洒至GSF设备，GSF设备根据GSE头信息转发至BORDER设备，BORDER设备解封还原成IP报文并转发至CORE设备，CORE根据路由信息转发至目的PoD的BORDER设备，BORDER设备将IP报文重新封装为GSE报文，经目标PoD内GSF设备转发至目的GSP设备，目的GSP设备解封还原成IP报文，转发至目的服务器网卡。

其他组网架构优化

考虑到实际业务开展及部署实施的差异化需求，GSE网络的组网架构可从多个不同方向进

GSE技术为十万卡级智算组网提供了无阻塞、高均衡的核心技术支撑，能够有效解决传统组网技术痛点，推动智算集群向更大规模演进，并为构建新一代智算网络、赋能智算产业高质量升级提供有力保障。



▲图3 GSE多PoD互联（GSP直出方式）组网示意图

行针对性优化。

- 收敛比配置：考虑到跨PoD流量通常小于PoD内流量，可按需针对性配置收敛比。其中，三层组网架构的收敛比部署在GSF-T2设备，多PoD互联架构的收敛比部署在BORDER设备。常规推荐收敛比为1:7。
- 单PoD规模最大：三层组网架构中，可在GSF-T2层采用轨道化部署。将原Full Mesh全互联方式，调整为不同SU单元中同序号的GSP设备与若干台GSF-T2设备互联，组成一个轨道；不同序号的GSP设备和其他GSF-T2设备互联，组成其他若干轨道。调整后单PoD最大可支持32K卡规模，且轨道化部署与Full Mesh部署构建十万卡级组网的设备及光模块总用量保持一致。
- PoD间跳数最少：多PoD互联架构中，为减少PoD间流量转发跳数，可将PoD内简化为GSP设备和GSF设备的两层组网，如图3所示，PoD间由GSP设备直接接至PoDI的CORE设备上，使跨PoD流量可直接接入PoDI网络，减少经过PoD内的跳数。该方式对于PoD内网络，可在GSP设备上到GSF方向部署加速比，提高网络吞吐；对于PoD间网络，可在GSP设备上到PoDI方向部署收敛比。

GSE技术为十万卡级智算组网提供了无阻塞、高均衡的核心技术支撑，能够有效解决传统组网技术痛点，推动智算集群向更大规模演进，并为构建新一代智算网络、赋能智算产业高质量升级提供有力保障。ZTE中兴

数据中心光模块的演进

袁智勇
中兴通讯数据中心交换机规划代表

在算力即生产力的时代，数据中心已不仅是传统的服务器机房，更是数字世界的“心脏”。而在这颗心脏中，长期被视作“神经末梢”的光模块，正悄然经历一场深刻变革。

过去十年间，光模块技术步伐相对平缓，从100G向400G的迭代已成为行业常态。然而，面对高密度算力需求，传统光模块在功耗、延迟与布线复杂度等方面的局限日益凸显，正逐渐成为制约数据中心网络能效的“最后一公里”瓶颈。

自2025年起，随着AI训练集群、大模型推理等高密度流量场景的爆发，AI算力竞赛正以前所未有的力度，驱动光模块技术加速演进。目前，800G/1.6T光模块已步入早期部署与测试阶段，其中功耗控制成为关键研发方向，预计将在2026年迎来规模化应用。与此同时，更前沿的3.2T光模块研发也已启动，亟待突破芯片与封装技术瓶颈，以支撑下一轮算力升级。

技术演进

当前在智算数据中心规模部署的51.2T交换机中，每台设备通常配备128个400G光模块。以典型的400G FR4光模块为例，其功耗约为10W，这使得单台交换机在全速运行下的总功耗接近3000W，其中光模块所占功耗比重已超过40%。

传统400G光模块普遍采用“电-光-电”三级转换架构：交换芯片发出的电信号，经由光模块内部的DSP芯片处理，再通过激光器转换为光信号进行传输，到达对端后重新转换为电信号。该架构稳定可靠，支撑了全球超过80%的400G应用部署，但也带来明显短板：功耗占比高、信号路径长、抖动累积显著，且进一步提升速率与传输距离的难度较大。

随着端口速率向1.6T演进，沿用传统架构的光模块功耗预计将突破20W，散热压力急剧上升，机柜功率密度逐渐接近物理极限。至3.2T阶段，模块功耗可能高达50W，传统风冷已难以满足散热需求，而液冷方案又因结构限制（如鼠笼影响冷板接触效率）导致散热效能下降。若迈向6.4T及以上速率，电信号速率将超过200G/通道，PCB损耗、连接器阻抗及金手指串扰等信号完整性挑战将愈发突出，解决成本与难度显著增加。

面对以上挑战，光模块技术在不断演进创新，出现了LPO线性可插拔光模块、LRO线性接收光模块、NPO近封装光学方案、CPO共封装光学等技术。LPO技术通过去除光模块中的数字信号处理器，大约能降低50%的功耗。LRO技术则是一种折中方案，仅去除发送方向的DSP，相应地功耗降低幅度也较小。NPO/CPO技术较为激进，可大幅降低功耗。

LPO线性直驱

LPO（linear pluggable optics）通过去除模块内的DSP芯片，将信号处理功能转移至交换芯片中，使光模块仅专注于光电转换。该技术主要

优势包括：端口功耗可降至5W（以400G为例），实现节能50%以上；因省去高价值DSP芯片，整体成本下降约15%~20%；同时支持向800G/1.6T的平滑演进，兼容现有光接口标准。

然而，LPO技术也具有一定局限性。它依赖交换芯片具备高精度线性驱动能力，TIA（transimpedance amplifier，跨阻放大器）和驱动芯片无法完全取代DSP，且由于信号处理简化，系统误码率相对较高，传输距离因此受限，目前主要适用于数据中心内部500m以下的短距离互联。此外，LPO当前主要适配400G/800G端口，更高速率的标准化与生态尚未成熟。在部署层面，LPO需与交换机ASIC芯片进行深度参数调优，系统调试较为复杂，且跨厂商互联互通仍有待充分验证，整体产业生态仍处于发展阶段。

LRO线性接收

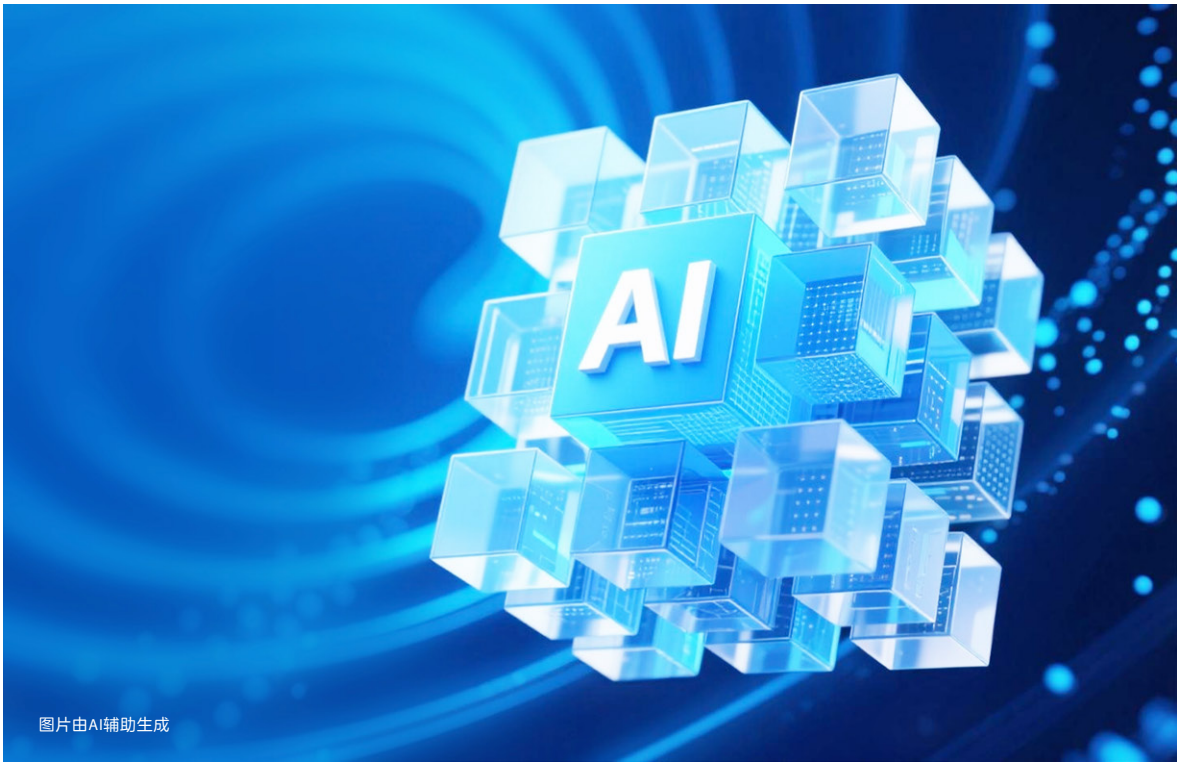
LPO技术通过完全移除DSP芯片来降低功耗与成本，也因此面临传输距离受限和系统互操作性方面的挑战。

相比之下，LRO（linear receive-side optics）作为一种折中方案，在发射端保留DSP芯片，在接收端采用线性设计，旨在兼顾性能与能效。其功耗低于全DSP传统光模块，传输距离优于LPO，但目前仍受限于规模效应不足、产业链成熟度较低，尚未成为市场主流选择。

NPO光电协同

NPO（near-package optics）可视为LPO的进阶形态。其核心设计是将光引擎从可插拔模块中剥离，并集成到交换芯片封装附近，通过高性能基板实现两者的近距离互连（通常间距<150mm，信道损耗≤13dB）。这种架构大幅缩短了电信号路径，在提升集成度的同时有效降低了信号衰减。

与当前主流的可插拔光模块相比，NPO的互联密度可提升2~3倍，代表了光互联向更高集成度演进的关键过渡阶段，也为后续CPO（共封装光学）技术的落地奠定了工程基础。此外，NPO



图片由AI辅助生成

技术对比	速率（单位：bps）	功耗	传输距离	成本	产业成熟度
传统技术	100G/200G/400G/800G，1.6T开始商用	高	>10km	高	高
LPO线性直驱	400G/800G	较低	<500m	较低	中
LRO线性接收	400G/800G	中	<500m	中	低
NPO光电协同	3.2T/6.4T，拆分支持多个400G/800G	低	<2km	低	低
CPO光电一体	3.2T/6.4T，拆分支持多个400G/800G	最低	<2km	最低	较低

▲表1 各光模块技术对比

结构支持光引擎与交换芯片解耦，有助于降低对单一芯片供应商的依赖，在供应链层面提供更多灵活性。

CPO光电一体

CPO（co-packaged optics）代表了光互联技术演进的终极方向：它将光引擎与电芯片集成封装在同一基板或中介层上，从而完全避免了传统PCB上的长距离电信号传输。其电信号传输距离通常低于50mm，信道损耗控制在7dB以内。

在能效方面，CPO表现突出。例如，3.2T CPO方案的功耗约为18W。对于一台51.2T交换机而言，仅需16个此类模块，相比传统可插拔光模块，整体功耗可降低77%以上。该技术不仅显著提升了互联带宽密度、降低了系统误码率，还能大幅节省交换机面板空间，有效突破前面板端口密度的物理限制。

然而，CPO也带来了新的可靠性与维护挑战。其光引擎依赖外部激光源（ELS）提供光信号，一旦该部件故障，将导致多个光端口同时失效，可能引发局部网络中断。此外，CPO中的光引擎与电芯片为共封装设计，无法像传统可插拔模块那样进行独立更换，任一组件故障都可能需更换整个封装体，因此对光引擎的良率、长期可靠性及维护策略提出了极高要求。

目前，CPO仍处于发展早期，整机级别的可靠性尚需在实际网络环境中进行长期验证。尽管如此，其在超高带宽、超低功耗及高密度互联方面的巨大潜力，已使其成为未来光通信——尤其是AI算力集群与超大规模数据中心——不可或缺的关键技术方向之一。

各光模块技术的对比见表1。

主要标准

相关电气接口标准主要由OIF（Optical Internetworking Forum，光互联网论坛）主导推进。当前各技术路径的标准成熟度存在差异，直接影响了其规模化部署的可行性：

- LPO标准已正式发布，生态相对清晰，为当前部署提供了明确依据。
- LRO的标准草案已发布，正处于完善阶段，

为后续产品兼容性奠定基础。

- CPO方面，3.2T的标准已发布，而更高密度的6.4T标准仍在制定中，技术路径尚未完全固化。

业界部署情况

目前，LPO技术已在头部云厂商和芯片企业（如阿里云、英伟达、Meta）的数据中心内部互联及AI训练集群中实现规模部署，有效降低了PUE和延迟，提升了能效与成本效益。而CPO技术尚处于实验室验证与规划阶段，Meta已完成基于博通方案的51.2T交换机可靠性验证，Google、Microsoft等云商计划于2026至2028年逐步规模化部署，国内则以试点探索为主。

LPO

头部云与芯片厂商已在关键场景中率先导入LPO技术，其应用路径与价值导向具有重要参考意义。

阿里云于2024年在其基础设施网络中规模部署LPO，主要用于数据中心内部互联。该方案显著降低了整体PUE，并为AI训练集群的扩展提供了高能效、低成本的连接支撑。

英伟达在其内部AI集群（如GB200 NVL72）中采用LPO实现GPU间高速互联，自2024年起进入量产应用阶段。英伟达特别强调该技术带来的低延迟优势，将其视为提升大规模AI训练效率的关键一环。

Meta2024年上半年开始导入LPO，用于RSC等AI训练集群的短距离互联，并结合硅光技术进一步优化整体能效表现，体现了其在追求算力密度与能效平衡方面的技术路线选择。

CPO

Meta在CPO技术的验证上取得了关键进展，基于博通Bailly的51.2T CPO交换机（集成8个6.4T硅光引擎）完成实验室验证，证明其具备高可靠性，为面向智算业务升级网络架构提供了明确的技术路径与可靠性参考。

与此同时，Google、Microsoft、Amazon等云服务商正在规划或试点CPO扩展方案，预计2026至2028年逐步转向规模化部署。

国内互联网公司也在智算中心开展CPO与全光融合方案的探索，但目前公开的商用案例仍以试点为主，规模化应用尚待进一步推进。

我们的思考

在当前光模块技术快速演进的背景下，LPO、LRO、NPO与CPO分别代表了不同阶段的技术选择，各有其适用场景与权衡点。

LPO通过去除模块内DSP，可实现约50%的功耗降低，成本效益显著，已在部分互联网企业的特定场景中开始试用。但其对交换芯片线性驱动能力要求较高，且在多厂商设备互联时仍面临兼容性与调试复杂性的挑战。

LRO作为折中方案，仅在接收端去除DSP，在兼容性和功耗之间取得更好平衡，虽节能幅度不及LPO，但系统适应性与部署难度相对较低。

NPO可视为技术演进中的过渡形态，为后续CPO的实现奠定集成基础，目前尚未形成规模部署。

CPO虽在理论上具备显著的能效与密度优势，但其高度集成的特性也带来了可维护性低、故障影响范围大、供应链依赖性强等运维层面的挑战，目前仍处于早期验证与试点阶段。

综合技术成熟度、部署灵活性及运维风险等因素，LPO可作为当前阶段具备可落地性的优先选项，尤其适用于对功耗敏感、距离较短且设备生态可控的场景。而CPO更适用于未来超高密度、超高能效要求的定向场景，需伴随标准、生态及可靠性经验的逐步完善，方可在规模化部署中发挥其潜力。ZTE中兴

铸就中部地区智能计算新基座

——中兴通讯助力湖南移动算力资源池网络建设

随着人工智能技术的迅猛发展，算力已成为驱动数字经济发展的核心动力。大模型训练与推理需求的激增，使传统计算基础设施面临严峻挑战。特别是在国家强调自主可控、信息安全的战略背景下，构建基于国产技术的算力资源池成为当务之急。

湖南省工业基础雄厚，拥有千亿级企业4家、百亿级企业53家，并已形成多个国家级先进制造业集群。当前，省内企业正迫切需要通过AI技术提升生产与管理效率，以应对市场竞争，推动产业升级。然而，企业应用AI技术普遍面临算力获取成本高、技术门槛高、数据安全保障等挑战。通过建设国产化算力资源池，提供集约化、规模化的普惠算力，可有效降低企业的AI技术使用门槛。

2025年12月，湖南移动国产算力资源池正式点亮投运。该项目是中国移动落实国家构建全国一体化算力体系战略、强化自主创新能力在区域层面的关键部署，标志着湖南移动在新型信息基础设施建设上取得重大进展。该资源池将聚焦区域发展需求，为长沙智能制造、智慧城市、数字文创、生物医药、智能网联汽车等重点产业与前沿领域的创新研发及应用落地，提供坚实的算力保障与强劲的创新动能。

秦芳
中兴通讯交换机产品策划经理

全栈国产化智算网络方案

在AI大模型参数量以每年10倍速度增长的背景下，智算能力需求呈指数级攀升。受限于单卡算力性能，需要部署更多GPU卡来提升整体算力性能，这使得大规模智算组网能力在智算业务建设中尤为重要。湖南移动算力资源池项目采用软硬件协同设计，构建了从底层硬件到上层应用的全栈国产化解决方案，打造出千卡规模智算集群。

其中，壁仞科技提供高性能国产GPU，为大规模计算提供强劲动力；中兴通讯星云智算网络解决方案，则成为整个资源池高效运行的关键支撑。

弹性极简架构，支撑大规模集群平滑扩展

湖南移动算力资源池网络的参数面采用框盒混合的Spine-Leaf两层胖树组网架构，实现带宽无收敛和轨道化设计，支持GPU卡200G接入及RoCE无损组网。Spine层和Leaf层采用中兴通讯ZXR10 9900X系列核心交换机和ZXR10 5960M系列数据中心交换机，实现网络带宽无阻塞。样本面同样采用Spine-Leaf两层胖树架构及带宽无收敛设计，服务器100G接入，RoCE无损组网。Spine和Leaf层均采用中兴通讯ZXR10 5960M系列数据中心交换机。该架构可依据不同阶段的智算建设规模需求，进行灵活弹性扩展，为未来规模增长预留空间。

超高性能实现，锻造高效无损网络

项目应用的中兴通讯ZXR10 9900X系列核心交换机采用业界领先的CELL/VOQ交换架构，在设备交换性能上全面超越传统Packet交换架构。中兴通讯还是国内首个完成112Gbps SerDes产品化的厂商，ZXR10 9900X系列核心交换机单机最大可支持576个400G/800G端口，是国产最高密度400G/800G框式交换机。

针对传统RoCE网络在拥塞控制精度方面的性能限制，项目采用中兴通讯创新的ENCC高精度端网协同拥塞控制技术，将端到端网络带宽利用率从60%大幅提升至98%以上。同时，为了解决传统负载均衡策略的局限性，采用中兴通讯层次化负载分担架构和智能全局负载分担方案，确保全网链路的负载分担效率最优。

自主可控，全栈国产化保障智算产业安全

项目采用全国产化技术架构，硬件层面基于全国产GPU、国产交换机，以中兴通讯自主研发的DPU芯片及大容量交换芯片为核心，通过业界领先的OLink开放互联平台，实现国产化GPU卡的大规模高性能互联；软件层面搭载自主研发的操作系统及配套软件体系，实现从底层硬件到上层应用的全栈国产化，充分保障了算力资源的国家信息安全与技术自主可控。

开放解耦，构建敏捷高效的智算新生态

项目从基础设施、能力平台、算力网络层面构建全栈开放的智算方案，并通过软硬解耦、训推解耦、模型解耦，推动各类能力组件化和共享赋能，加速AI技术的创新、研发、应用的商业化进程。

项目构建的“算力+平台+生态”一体化模式，为湖南省及周边区域的智能制造、智慧城市等优势产业提供强大算力支撑。该模式不仅降低了企业使用先进算力的门槛，还推动了人工智能技术在传统产业的渗透与应用。

网络创新带来的多重收益

湖南移动国产算力资源池的建设，通过网络技术的创新突破，为区域数字经济发展提供多重价值。

技术性能实现跨越式提升。项目采用的创新网络方案将网络吞吐效率提升至98%以上，拥塞调整时间降至微秒级别，大幅提升了算力利用效率。千卡GPU集群断点续训平均恢复时间小于5分钟，保证了大规模型训练任务的稳定性与连续性。

区域产业升级获得新动能。湖南移动表示，将搭建产业合作平台，整合产业链上下游资源，构建研发、转化、应用闭环生态。同时，湖南移动将面向高端装备制造、汽车电子信息等重点产业，提供定制化专享智算解决方案，助力“湖南制造”提质升级。

湖南移动国产算力资源池项目的成功实践表明，国产化技术路线不仅能满足国家信息安全战略需求，还能在性能上达到行业领先水平。随着技术不断迭代，这一高效、稳定的国产化智算网络将成为推动中部地区数字经济与实体经济深度融合的核心引擎，为区域高质量发展注入持续动能。ZTE中兴

中兴通讯助力河南移动、广西移动打造800G以太网跨域智算互联新标杆

随着人工智能、云计算等技术的蓬勃发展，特别是随着DeepSeek大模型等AI大模型的广泛应用，互联网流量呈现爆发式增长，传统网络架构在应对大规模分布式训练、实时数据交互等场景时面临严峻挑战。智算中心之间的高效互联，尤其是跨地域、跨楼宇的数据传输，对网络带宽、时延提出了更高要求。

中兴通讯凭借业界领先的超高速800G以太网互联技术，联合河南移动、广西移动完成国内核心路由器800G以太网跨域高速互联方案现网试点，打造跨域智算互联网络新底座。该方案围绕传输性能、单板转发能力、业务处理能力三大核心维度展开试点验证，结果显示各项指标均达预期。配备800GE端口的ZXR10 T8000-X16核心路由器，可全面兼容Native IP与组播流量转发、L3EVPN over SRv6 Policy以及等价/非等价负载均衡等功能。中国移动客户表示：“这一成果强有力地证实800G以太网技术在应对未来网络对更高带宽互联需求方面的出色性能与可靠性”。该方案不仅突破了传统100G/400G网络的带宽限制，也大幅降低数据往返时间RTT（round-trip time），使分布在不同地理位置的GPU集群能够实现近乎“零延迟”的协同计算。同时，该方案采用高密度、低功耗的硬件设计，在提升算力网络性能的同时，显著降低了运营成本，契合国家“双碳”战略下的绿色数据中心建设要求。

加速算力，决胜毫秒

智算中心跨域互联的核心目标是通过高速、低时延的网络，将分布于不同地域的智算中心互联，实现跨区域的大规模数据计算和处理能力。河南移动、广西移动采用的中兴通讯800G以太网

黄霜霜
中兴通讯有线产品规划总监

刘义
河南移动规划技术部数据通信资深专家

超高速端口，可提供8倍于100GE、2倍于400GE的互联带宽，大幅提升数据搬运效率，突破智算时代的“互联瓶颈”。

带宽方面，传统网络架构下，GPU集群间的数据交换往往受限于带宽，导致训练效率下降，而800G技术将网络从“通道”升级为“加速引擎”，确保海量参数与训练数据能在算力单元间无阻塞地自由流动。

时延方面，AI训练对时延极为敏感，尤其是分布式训练场景下，数据往返时间RTT直接影响模型收敛速度。河南移动、广西移动采用的超高速800G以太网互联技术，将数据往返时间RTT压缩至毫秒级，使得跨地域GPU集群能够如同单机般紧密协同，显著提升训练效率。

随着AI模型参数规模突破万亿级，800G网络将成为支撑下一代智算中心的关键基础设施。河南移动、广西移动的试点验证不仅满足当前需求，更为未来3—5年的算力增长预留扩展空间。

降本增效，生态开放

传统智算中心互联依赖多个100G或400G链路捆绑，但负载不均导致实际带宽利用率受限。中兴通讯采用超高速800G以太网互联技术，显著提升整体传输效率和带宽。

降本方面，800G技术以其单端口高带宽特性，有效减少所需互联链路数量，降低对光纤与波长资源的占用。通过部署800G以太网互联，网络带宽利用率得到大幅提升，进而摊薄了光纤、设备及机房等基础设施的单位比特传输成本，实现整体成本优化。

增效方面，800G超高速以太网通过消除多链路捆绑带来的负载不均衡问题，减少带宽浪费，使网络规划更精准，运维复杂度显著降低。这不仅提升了网络运营效率，也为网络的长期稳定性和高可用性提供了支撑。

生态方面，超高速800G以太网互联技术基于开放标准，支持与业界主流设备无缝对接，避免厂商锁定问题。河南移动、广西移动可灵活引入第三方设备，进一步降低总拥有成本（total cost of ownership，TCO）。

绿色集约，永续发展

在全球“双碳”目标下，绿色节能成为智算中心的核心竞争力。800G跨域智算互联方案通过高集成度与能效优化，帮助河南移动、广西移动实现显著的节能减排。

能效方面，中兴通讯大容量核心路由器ZXR10 T8000-X16提供高密度800GE端口，单槽处理能力达14.4Tbps。相比传统400G方案，单板Gbit能耗降低23%，相比传统100G方案，单板Gbit能耗降低38%，有效减少碳排放，助力绿色数据中心发展。

空间方面，800G的高端口密度可减少50%以上的机柜占用，释放的宝贵空间可用于部署更多算力服务器，实现数据中心每平方米产值的最大化。

在可持续发展方面，打造绿色低碳的先进智算中心，既符合国家战略与全球发展趋势，也是河南移动、广西移动践行社会责任的重要体现。

中兴通讯与河南移动、广西移动合作完成的800G以太网跨域智算互联方案现网试点，标志着我国在超高速互联技术领域取得重要突破，正式迈入智算网络800G时代。这一创新方案通过超高速互联、生态开发和绿色节能三大技术优势，不仅为智算业务发展提供了强有力的网络支撑，也有效解决了当前AI算力瓶颈，更在全国范围内树立了算力网络升级的新标杆。随着AI算力需求的持续增长和“东数西算”工程的深入推进，800G技术将成为构建全国一体化算力网络的关键基石，在支撑数字经济高质量发展的同时，助力我国在全球科技竞争中占据领先地位，为数字中国建设提供坚实的技术底座。ZTE中兴

ZTE中兴

驭风 系列

纤薄至简 驭风随行

驭风10 Air

13.9mm厚度 | 1.25kg净重 | 14英寸FHD高清显示屏 | 5W超低功耗

丰富接口 | 全金属机身 | 无风扇设计

ZTE中兴

致力于成为网络连接和智能算力的领导者
让沟通与信任无处不在