



中文核心期刊 中国科技核心期刊 中国核心学术期刊  
第三届国家期刊奖百种重点期刊 信息通信领域产学研合作特色期刊

ISSN 1009-6868  
CN 34-1228/TN

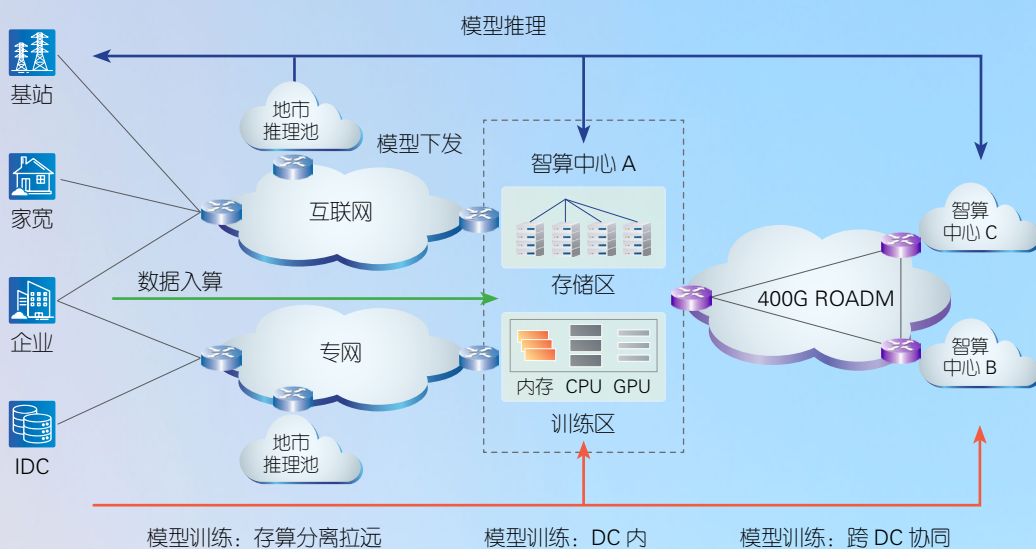
# 中兴通讯技术

## ZTE TECHNOLOGY JOURNAL

<http://tech.zte.com.cn>

第 31 卷 · 总第 182 期 · 2025 年 4 月 · 第 2 期

### 专题：智算网络演进及关键技术



(封面图片详解见 P32)

ISSN 1009-6868



第二十七届中国科协年会学术论文专刊

# 《中兴通讯技术》第10届编辑委员会

**顾问** 侯为贵(中兴通讯股份有限公司创始人) 钟义信(北京邮电大学教授)

陈锡生(南京邮电大学教授) 糜正琨(南京邮电大学教授)

**主任** 陆建华(中国科学院院士)

**副主任** 方 榕(中兴通讯股份有限公司董事长) 李建东(西安电子科技大学教授)

## 编委

陈建平	上海交通大学教授	唐雄燕	中国联通研究院副院长
陈前斌	重庆邮电大学教授、副校长	陶小峰	北京邮电大学教授
段晓东	中国移动研究院副院长	汪烈军	新疆大学教授、副校长
方 榕	中兴通讯股份有限公司董事长	王 翔	中兴通讯股份有限公司高级副总裁
葛建华	西安电子科技大学教授	王文博	北京邮电大学教授
管海兵	上海交通大学教授、副校长	王文东	北京邮电大学教授
郭 庆	哈尔滨工业大学教授	王喜瑜	中兴通讯股份有限公司执行副总裁
洪 伟	东南大学教授	王耀南	中国工程院院士、湖南大学教授
江 涛	华中科技大学教授	王志勤	中国信息通信研究院副院长
蒋林涛	中国信息通信研究院科技委主任	卫 国	中国科学技术大学教授
金 石	东南大学教授、副校长	邬贺铨	中国工程院院士
李尔平	浙江大学教授	吴春明	浙江大学教授
李红滨	北京大学教授	向际鹰	中兴通讯股份有限公司首席科学家
李厚强	中国科学技术大学教授	肖 甫	南京邮电大学教授、副校长
李建东	西安电子科技大学教授	解冲锋	中国电信研究院教授级高工
李乐民	中国工程院院士、电子科技大学教授	徐安士	北京大学教授
李融林	华南理工大学教授	徐子阳	中兴通讯股份有限公司总裁
林晓东	中兴通讯股份有限公司副总裁	续合元	中国信息通信研究院首席专家
刘 健	中兴通讯股份有限公司高级副总裁	薛向阳	复旦大学教授
刘建伟	北京航空航天大学教授	杨义先	北京邮电大学教授
隆克平	北京科技大学教授	易芝玲	中国移动研究院首席科学家
卢光跃	西安邮电大学教授、校长	张 杰	北京邮电大学教授
陆建华	中国科学院院士、清华大学教授	张 平	中国工程院院士、北京邮电大学教授
马建国	中原工学院教授、学术副校长	张 卫	复旦大学教授
毛军发	中国科学院院士、深圳大学校长	张宏科	中国工程院院士、北京交通大学教授
孟洛明	北京邮电大学教授	张钦宇	哈尔滨工业大学(深圳)教授、副校长
尼玛扎西	中国工程院院士、西藏大学教授	张云勇	中国联通云南分公司总经理
石光明	鹏城实验室副主任	赵慧玲	工业和信息化部信息通信科技委常委
史振威	内蒙古大学教授	郑纬民	中国工程院院士、清华大学教授
孙知信	南京邮电大学教授	钟章队	北京交通大学教授
谈振辉	北京交通大学教授	周 亮	南京邮电大学教授、副校长
唐 宏	中国电信 IP 领域首席专家	朱近康	中国科学技术大学教授
唐万斌	电子科技大学教授	祝宁华	中国科学院院士、南开大学教授

# 目次

中兴通讯技术 (ZHONGXING TONGXUN JISHU)  
第 31 卷 总第 182 期 2025 年 4 月 第 2 期

中文核心期刊 中国科技核心期刊 第三届国家期刊奖百种重点期刊 信息通信领域产学研合作特色期刊 中国知网、万方数据、重庆维普等数据库收录期刊 1995 年创刊

热点专题 ▶	智算网络演进及关键技术
	01 专题导读 ..... 段晓东, 李丹, 虞红芳
	03 星脉网络:面向 GPU 集群集合通信与集中式路由的协同优化 ..... ..... 李宝嘉, 何春志, 夏寅贲, 何泽坤, 王晓亮
	14 智算中心 Scale-Out 网络的演进及 GSE 的实践 ..... 程伟强, 李新双, 白艳, 吕勇
	21 基于大语言模型蒙特卡洛树搜索的智算网络故障根因分析系统 ..... 罗子秋, 苗宇锴, 李丹
	31 算力智联网关键技术研究 ..... 易昕昕, 张乃瞻, 刘雅承, 韩梦瑶, 曹畅
	39 面向多算力中心协同的广域智算网络仿真架构设计 ..... 边彦晖, 刘明远, 虞红芳
	47 在网计算:愿景与关键技术 ..... 刘杨, 孙浩南, 程伟强
名家视点 ▶	53 智算网络发展综述 ..... 段晓东, 程伟强, 张昊
企业视界 ▶	63 智算中心组网技术及应用 ..... 段威, 于浩, 李和松, 张征, 王宁诚
技术广角 ▶	72 面向 5G 网络的工业互联网安全技术分析 ..... 陈焱, 陈丹, 袁琦, 刘小丽, 徐晓娜
	77 物理层安全保密电话的实现 ..... 林立峰, 周子健, 焦秉立

## 《中兴通讯技术》2025 年热点专题名称及策划人

### 1. 6G 立体覆盖技术

西安电子科技大学教授 李建东  
西安电子科技大学教授 刘俊宇

### 2. 智算网络演进及关键技术

中国移动研究院副院长 段晓东  
清华大学教授 李丹  
电子科技大学教授 虞红芳

### 3. 6G 网络安全

北京航空航天大学教授 刘建伟  
北京航空航天大学教授 王景璟

### 4. 面向 6G 的高时效智能机器通信技术

中国工程院院士、北京邮电大学教授 张平  
北京邮电大学副教授 秦晓琦

### 5. 网络中的 AI 技术

中国电信研究院教授级高工 解冲锋  
北京邮电大学教授 孟洛明  
清华大学教授 崔勇

### 6. 新一代光传输技术

上海交通大学教授 陈建平  
中国联通研究院副院长 唐雄燕

# MAIN CONTENTS

ZTE TECHNOLOGY JOURNAL  
Vol. 31 No. 2 Apr. 2025

Special Topic ►	<b>Evolution and Key Technologies of Intelligent Computing Networks</b>
	01 Editorial ..... DUAN Xiaodong, LI Dan, YU Hongfang
	03 Astral Network: Co-Optimizing Collectivce Communication and Central Traffic Routing for GPU Clusters ..... LI Baojia, HE Chunzhi, XIA Yinben, HE Zekun, WANG Xiaoliang
	14 Evolution of Scale-Out Network in Intelligent Computing Centers and Practice of GSE ..... CHENG Weiqiang, LI Xinshuang, BAI Yan, LYU Yong
	21 Fault Root Cause Analysis System of Intelligent Computing Networks Based on Large Language Models and Monte Carlo Tree Search ..... LUO Ziqiu, MIAO Yukai, LI Dan
	31 Key Technologies for Intelligent Computing Power Network ..... YI Xinxin, ZHANG Naihan, LIU Yacheng, HAN Mengyao, CAO Chang
	39 Wide-Area Intelligent Computing Network Simulation Architecture for Multi-Computing Centre Collaboration ..... BIAN Yanhui, LIU Mingyuan, YU Hongfang
	47 In-Network Computing: Vision and Key Technologies ..... LIU Yang, SUN Haonan, CHENG Weiqiang
Expert View ►	53 Development of Intelligent Computing Networks: A Survey ..... DUAN Xiaodong, CHENG Weiqiang, ZHANG Hao
Enterprise View ►	63 Intelligent Computing Center Networking Technology and Applications ..... DUAN Wei, YU Hao, LI Hesong, ZHANG Zheng, WANG Ningcheng
Research Papers ►	72 Analysis of Industrial Internet Security Technology and Development for 5G Network ..... CHEN Yan, CHEN Dan, YUAN Qi, LIU Xiaoli, XU Xiaona
	77 Realization of Wireline Telephone Based on Physical Layer Security ..... LIN Lifeng, ZHOU Zijian, JIAO Bingli

期刊基本参数: CN 34-1228/TN\*1995\*b\*16\*82\*zh\*P\*¥20.00\*6500\*11\* 2025-04

敬告读者	本刊享有发表文章的版权, 包括英文版、电子版、网络版和优先数字出版版权, 所支付的稿酬已经包含上述各版本的费用。未经本刊许可, 不得以任何形式全文转载本刊内容; 如部分引用本刊内容, 须注明该内容出自本刊。
------	---



# 智算网络演进及关键技术 专题导读



## 专题策划人



段晓东



李丹



虞红芳

大模型技术已从探索阶段迈向普惠应用，呈现爆发式增长与深度行业渗透的态势。当前人工智能（AI）技术发展呈现双重趋势：一方面遵循传统扩展定律（Scaling Law），通过“大模型+大算力+大数据”持续突破模型性能边界；另一方面推进“模型算法+算网设施”协同优化，以系统级创新探索算效极限，构建高性价比解决方案。二者的同步推进正驱动全球算力基础设施呈指数级增长。

智算互联网络作为大规模图形处理器（GPU）集群分布式计算的核心基础设施，正面临架构升级与性能突破的双重挑战：一方面，大模型迭代速度的指数级增长推动算力集群规模从万卡级向十万卡量级跨越，驱动超大规模组网架构创新；另一方面，分布式计算效率对网络性能提出严苛要求，需构建超低时延、超高带宽、超高可靠的无损网络环境。这一量级跃升与性能需求，亟需智算网络实现关键技术突破。

智算网络可划分为3个层级架构：服务器/超节点内的GPU卡间互联网络、园区级智算中心机间互联网络、广域级智算中心间互联网络。各层级在拓扑结构、性能指标及技术

方案上存在显著差异：1）服务器/超节点内GPU互联网络需承载张量并行（TP）、序列并行（SP）、混合专家（MoE）并行计算任务，具备纳秒级延迟与太比特每秒级的带宽需求。国际先进方案已经实现256卡全互连架构，单卡吞吐量达900 Gbit/s；相较之下，中国GPU产业在互联规模与传输性能方面仍存在明显差距。2）智算中心内服务器互联网络支撑流水线并行（PP）、数据并行（DP）、MoE并行训练，需满足万卡至十万卡级无收敛组网需求，具备拍比特每秒级的交换能力与微秒级延时稳定性。该场景对交换芯片容量、无损传输及智能运维提出新要求。中国移动主导的全调度以太网（GSE）联盟基于以太网革新，构建标准开放的技术体系，产业生态逐步成熟。3）跨园区智算中心间互联网络主要处理PP、DP通信任务，面临带宽收敛、长距时延及数据安全三重挑战，需依托上层平台、算法模型与网络技术的协同优化实现突破。当前行业正加速推进长距跨智算中心联合训练的技术验证与标准体系构建，着力探索分布式智能计算的协同创新路径。

智算网络面临的核心挑战需通过算力与网络的深度融合实现系统性突破，其破局点涵盖架构设计、协议优化、故障恢复、芯片设计及生态协同等多维创新。本期专题就智算中心网络的发展趋势及挑战、超节点GPU互联协议、GPU集

群机间互联技术、多算力中心协同的广域智算网络保障和仿真、基于AI的智能运维等关键技术开展讨论。专题汇集了来自清华大学、电子科技大学、北京邮电大学、南京大学、中国移动、中国联通、腾讯以及中兴通讯等专家的最新研究成果，凝聚了他们多年的研究积累与实践经验。在此，谨向所有作者致以诚挚谢意。期待本专题能为读者提供有价值的学术参考和实践启示。

#### 策 划 人 简 介

**段晓东**，中国移动通信有限公司研究院副院长，教授级高工，享受国务院特殊津贴专家，工业和信息化部通信科技委信息通信专家组专家，“新世纪百千万人才工程”国家级人选；长期从事AI智算网络、下一代互联网、算力网络、5G网络架构、SDN/NFV下一代网络及6G网络等研究工作；先后主持了多项国家重大专项课题，并获得国家科技进步奖特等奖等多个奖项；获授权专利50余项。

**李丹**，清华大学教授、博士生导师，教育部长江学者特聘教授，IEEE Fellow，北京高校卓越青年科学家计划项目负责人，国家“973”计划项目首席科学家，国家重点研发计划项目负责人，国家“十四五”重点研发计划“网络空间安全治理”专家组副组长；主要从事计算机网络领域的研究工作；曾获教育部“青年科学奖”，以第一完成人获中国通信学会技术发明奖一等奖、中国电子学会技术发明奖一等奖；发表论文100余篇，获授权专利50余项。

**虞红芳**，电子科技大学教授、博士生导师、信息与通信工程学院副院长；长期致力于智慧网络及应用研究；主持多个国家自然科学基金重大项目、国家重点研发计划项目等国家级项目；曾获教育部自然科学奖二等奖、电子学会科技进步奖二等奖、四川省教学成果二等奖等奖项，主持研发的“跨数据中心高性能分布式机器学习系统GeoMX”和“基于轻量级虚拟化的大规模网络创新平台Klonet”分别获中国通信学会2021年未来网络领先创新科技成果奖、2021年网络5.0创新科技成果奖；发表论文100余篇，获授权中国发明专利30余项、美国发明专利2项，出版学术专著5本。

# 星脉网络:面向 GPU 集群集合通信与集中式路由的协同优化



## Astral Network: Co-Optimizing Collective Communication and Central Traffic Routing for GPU Clusters

李宝嘉/LI Baojia<sup>1</sup>, 何春志/HE Chunzhi<sup>1</sup>,  
夏寅贵/XIA Yinben<sup>1</sup>, 何泽坤/HE Zekun<sup>1</sup>,  
王晓亮/WANG Xiaoliang<sup>2</sup>

(1. 腾讯科技(深圳)有限公司, 中国 深圳 518000;

2. 南京大学, 中国 南京 210023)

(1. Tencent Technology (Shenzhen) Co., Ltd., Shenzhen 518000, China;

2. Nanjing University, Nanjing 210023, China)

DOI: 10.12142/ZTETJ.202502002

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250424.0957.004>

网络出版日期: 2025-04-24

收稿日期: 2025-03-24

**摘要:** 图形处理器 (GPU) 集群网络流量不断增加, 运营难度明显加大, 这给高性能大规模 GPU 集群网络系统的构建带来新的挑战与机遇。提出了一种能够实现超 10 万 GPU 集群互联的无损高性能网络方案——星脉网络。GPU 集群网络需要联合优化端侧的集合通信库和网络路由控制器, 以实现多路径的高效集合通信。为此, 针对星脉网络研发了端侧集合通信库 (TCCL) 以实现最短的跨节点路径规划, 同时还开发了全局优化路由器 (GOR) 以避免路径冲突导致的网络拥塞。在腾讯大模型 GPU 集群中, 星脉网络方案和公开 GPU 集群方案 (NVIDIA NCCL) 的对比结果表明: 星脉网络可以实现 25% 的集合通信带宽提升, 同时避免 80% 的由流量冲突造成的网络拥塞问题。

**关键词:** 大规模 GPU 集群; 集合通信; 负载均衡

**Abstract:** The network traffic of the graphics processing unit (GPU) cluster is continuously increasing, and the operation complexity has significantly increased, which brings new challenges and opportunities to the construction of high-performance large-scale GPU cluster network systems. To address this, we propose Astral Network—a lossless high-performance network architecture capable of interconnecting over 100 000 GPUs. GPU-centric networks require joint optimization of the collective communication library at the host and centralized routing controller to achieve efficient collective communication over multiple paths. Therefore, Tencent developed a collective communication library (TCCL) for Astral Network to achieve the shortest path planning across nodes, and a global optimized router (GOR) to avoid network congestion caused by route conflict through traffic planning. In Tencent's large-scale GPU-centric clusters, the comparison results between Astral Network and the publicly available GPU-centric network (i.e., NVIDIA NCCL) show that: Astral Network achieves 25% higher collective communication bandwidth while reducing traffic conflict-induced congestion by 80%.

**Keywords:** large-scale GPU clusters; collective communication; load balancing

**引用格式:** 李宝嘉, 何春志, 夏寅贵, 等. 星脉网络: 面向 GPU 集群集合通信与集中式路由的协同优化 [J]. 中兴通讯技术, 2025, 31(2): 3–13. DOI: 10.12142/ZTETJ.202502002

**Citation:** LI B J, HE C Z, XIA Y B, et al. Astral network: co-optimizing collective communication and central traffic routing for GPU clusters [J]. ZTE technology journal, 2025, 31(2): 3–13. DOI: 10.12142/ZTETJ.202502002

生成式人工智能 (AIGC) 的迅猛发展, 推动着人工智能 (AI) 大模型参数量从亿级飙升到万亿级<sup>[1]</sup>。模型参数规模增长与架构升级对底层网络提出新的要求。一个拥有 4 000 张图形处理器 (GPU) 卡的集群, 包含超过 150 个网络设备和 1 万条路由路径。若要支持 1.6 万张 GPU 卡互联, 网络设备数量需要超过 1 000 个, 路由表项超过 5 万条。集群规模越大, 所产生的通信损耗就会越高<sup>[2]</sup>。同时, AI 训练的通信模式与传统的通信模式差异较大, 不同大模型架

构在通信模式上也各有不同。部分大模型训练过程中, 通信占比最高可达 50%。分布式计算模式意味着, 单点故障可能导致整个集群不可用。因此, 在故障发生时, 需要快速定位问题并恢复训练, 将损失降到最低。如何在大规模组网的背景下, 提升通信效率, 降低通信占比, 保障训练的稳定性与高可用性, 进而提升 GPU 利用率和模型训练效率, 是 AI 网络亟待解决的核心问题。本文介绍了腾讯星脉网络应对这些挑战的思路和解决方案。

星脉网络架构设计基于两个原则：可扩展性和高可靠性。可扩展性是指能够根据AI业务训练的需求动态扩展集群规模，保证集群通信性能呈线性提升。可靠性是指保证架构能够承受一定程度的网络设备故障风险，使AI业务训练不出现中断。为符合上述原则，我们采用图1所示的多轨道网络架构。该架构通过同轨道交换机连接每个服务器中具有相同序号的网卡，然后通过二层（Spine）交换机实现一层（Leaf）交换机的全互联，最终形成一个两层的多轨道网络拓扑。具体地，为了实现跨GPU服务器的高速互联，每个服务器均配有8个端口速率为400 Gbit/s且支持基于融合以太网远端内存直接访问技术（RoCE）的网卡（NIC）。每个网卡通过同轨道的Leaf层交换机与其他服务器的同序号网卡实现互联。如图1所示，第一台服务器的NIC1到第二台服务器的NIC1的流量通过Rail 1交换机传输。此外，网卡采用双端口配置，每个端口连接到不同交换机。这可以有效降低端口故障的影响，提升集群互联的可靠性。网卡采用RoCEv2协议，并且使用优先级流量控制（PFC）实现无损网络。为了有效避免流量拥塞，我们采用带有动态水线的数据中心拥塞控制机制（DCQCN），以保证高通信吞吐。

## 1 LLM训练对GPU集群网络带来的挑战

以大语言模型（LLM）训练为主的GPU集群与传统数据中心或高性能计算场景有显著不同。深入理解和利用这些特性对于优化训练过程至关重要。LLM训练网络的特点主要表现在以下3个方面：

1) 网络边界扩展到服务器内。传统网络规划主要关注节点间互联，而GPU集群网络需同时考虑节点内NVLink/NVSwitch与节点之间RoCE连接的异构互联架构。这种扩展

的网络边界要求通信算法进行针对性设计：一是，突破传统NIC间传输范式，实现跨节点GPU显存之间的数据交换；二是，针对异构互联架构，开发多路径带宽聚合方案。

2) GPU通信模式呈现稀疏性和周期性。如图2（a）所示，以往大规模RoCE网络设计，基本基于两种假设：一是网络内所有节点可任意互联；二是流量随机且不可预测。图2（b）展示了我们从大模型训练网络中收集的NIC出口流量数据在每个训练迭代后，所呈现出的周期性和可预测性特点。这种结构化的流量模式为规划路由时最大化GPU集群的利用率提供了参考。

3) 大模型训练对网络抖动敏感。尽管传统网络在每个链路上承载大量流，但流的数量多且大多数流较小。在这种情况下，链路负载相对均衡。相对而言，GPU网络中的流数量相对较少，但每个流的体积却较大。例如，在一个拥有1 000个NVIDIA A100卡的集群中，每个流的峰值带宽可达180 Gbit/s，而此时流的数量仅为4 000，这容易导致网络拥塞。交换机不能单纯依靠与终端主机拥塞控制的交互来平衡流量，否则可能导致吞吐量下降。因此，我们需要一种有效的流负载均衡策略和动态调度机制，以避免拥塞状况，同时充分挖掘网络多路径的潜力。

为应对上述挑战，我们提出了一种新型流量管理框架。该框架借助集合通信优化和集中式路由规划的协同设计，实现以下关键成效：集合通信时延降低25%，链路带宽利用率提升至90%，网络拥塞发生概率控制在1%以内。

## 2 星脉网络设计概述

### 2.1 系统架构

星脉网络系统架构如图3所示，它由两个关键模块组

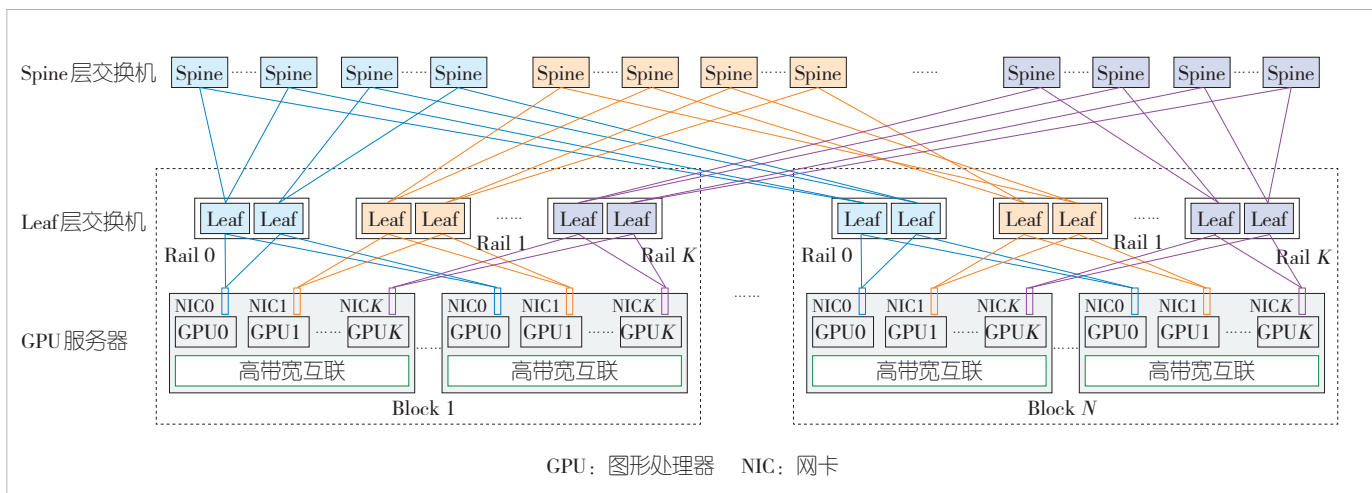


图1 GPU集群的多轨道网络架构



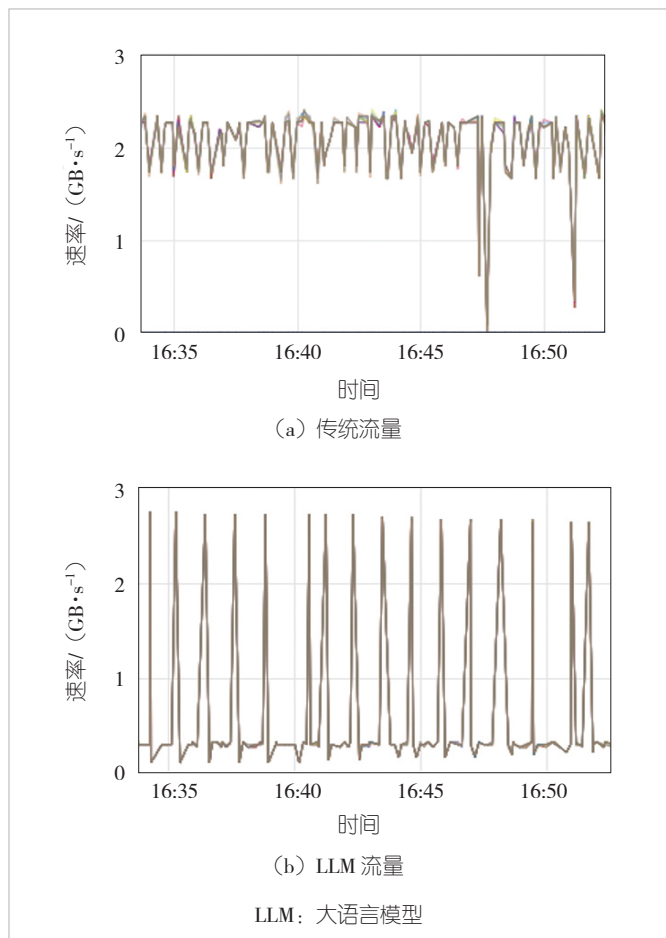


图2 从网卡出口观察到的传统流量与LLM流量

成：运行在GPU服务器上的拓扑感知集合通信库（TCCL）和管理RoCE网络中流量路由的全局优化路由器（GOR）。TCCL在GOR控制器的帮助下实现服务器内和服务器的整体流量预规划，GOR控制器则提供拓扑信息并指导RoCE网

络中的流量负载均衡。此外，GOR还负责在实际运营网络中的网络流量动态调度：首先通过网络流量和显式拥塞控制信号（ECN）来定位网络中拥塞链路，然后通过监控定位网络中的拥塞，最后通过交换机哈希注入的方式调整流量分配策略，以解决流量拥塞问题。

## 2.2 拓扑感知的集合通信库

TCCL针对大规模GPU集群通信场景，在标准英伟达集合通信库（NCCL）基础上实现3项核心优化：基于多级拓扑感知的通信路径优化算法、NVLink与RoCE网络的混合传输调度机制、确定性路由策略驱动的无冲突数据传输。与传统方案相比，本设计通过引入机内和机间拓扑数据感知模块，可有效弥补现有集合通信库在GPU集群网络架构感知能力上的缺陷。

当前主流集合通信库（如NCCL）依赖默认的服务端顺序决定集合通信路径规划。其构建典型通信路径的方式，遵循先机内通信后机间通信的原则。例如，服务器内的传输路径为：NIC0→GPU0→GPU1→……→GPU7→NIC7。跨服务器通信路径则是将每个服务器内的传输路径首尾互联，即Server1（NIC0→GPU0→……→GPU7→NIC7）→Server2（NIC0→GPU0→……→NIC7）→Server3（NIC0→GPU0→……→NIC7）。这种设计导致两个关键问题：一是相邻服务器间的NIC7→NIC0会产生跨轨道流量，这会增加网络拥塞概率；二是在多Block拓扑场景下（如Server1/3同属Block A，Server2位于Block B），默认跨机连接顺序会导致频繁的跨Block通信流量，造成二层网络中流量拥塞概率提升50%。为此，TCCL引入拓扑感知模块，通过解析服务器互连拓扑以及Block划分等特征，动态构建符合最小跳数原则

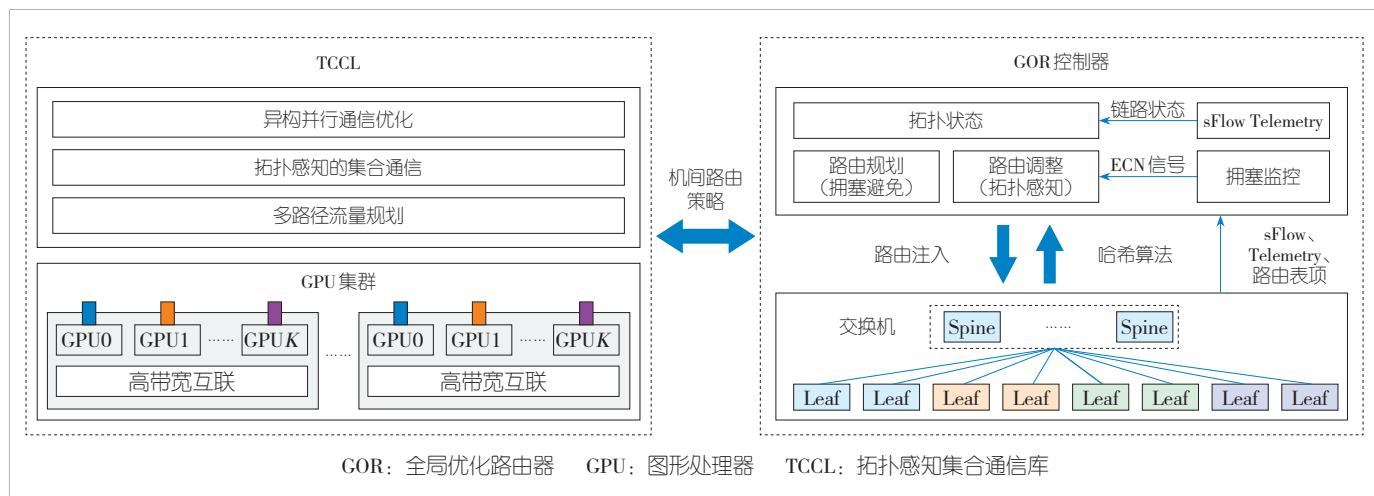


图3 星脉网络结构

和轨道亲和性的通信路径。

此外,针对异构网络资源的协同利用问题,TCCL还引入了混合通信传输机制。由于机内互连网络NVLink(通信带宽为400 GB/s)和机间互连网络RoCE(400 Gbit/s × 8 NIC = 3.2 Tbit/s)的通信带宽与通信协议有较大差异,TCCL引入了一种动态滑动窗口机制。该机制根据不同网络上的通信吞吐能力决定每个网络上承载的数据量,以最大化异构网络的带宽利用率。

为解决RoCE网络的哈希冲突问题,TCCL实现了确定性队列对(QP)连接管理机制。相较于NCCL随机生成QP连接参数的策略,本方案通过GOR控制器实时采集网络设备的等价多路径路由(ECMP)的哈希策略(包括源/目的IP、端口号等哈希因子),据此构建全局流表项冲突预测模型。为每个QP连接动态分配哈希正交的源端口号,可确保任意两条并发流量的哈希路径不会发生哈希冲突。

### 2.3 全局优化路由器

网络动态抖动会影响负载均衡,主要体现在以下4个方面:1)预规划通常根据哈希规则将流的数量均匀分布在网络链路上,但对于不同大小的流,局部拥塞仍可能发生。2)在大规模集群中,网络链路、交换机端口和GPU的故障是不可避免的,且出现故障的概率会随着设备数量的增加而增加。设备故障会导致拓扑不对称,减少可用路径,并使ECMP策略下的下一跳路径发生变化。这种路径重新分配会引发流量冲突。3)在多个模型同时运行的场景中,网络流量分布会随任务的启停时间动态变化。由于静态流量规划无法适应这种实时波动,因此在动态场景下容易出现资源竞争和流量冲突问题。4)在生产环境中,部分用户使用自定义的集合通信库(如私有消息传递接口实现),由于此类库并未纳入网络流量规划流程,可能导致协议协商失败,进而引发流量冲突。

为此,我们构建了一个基于集中控制器的动态流量管理系统。集中控制器负责实时监控网络状态,及时检测故障点并识别导致拥塞的流。控制器使用sFlow和遥测工具主动收集网络状态,并在模拟器中运用交换机厂商的实际ECMP哈希算法,以便为路由调度提供指引。当接收到拥塞信号时,全局控制器将受影响的流以最优方式调度到负载最低的路径上。

## 3 静态流量路由规划

### 3.1 LLM训练的流量特征

3D并行技术涵盖张量并行(TP)、流水线并行(PP)和数据并行(DP)技术,在LLM大模型的训练中得到广泛应

用<sup>[3]</sup>。DP借助在不同GPU组上部署多个模型参数副本的方式,将数据集划分为小批量,分发给不同GPU组进行并行训练。PP和TP将单个GPU无法容纳的大模型切分到多个设备上。

在整个流程中,每个阶段需执行两次TP AllReduce操作,分别用于Attention层和多层感知机(MLP)层的前向与反向传播,使得通信频率显著高于其他集合通信类型。此外,TP AllReduce的后续计算必须依赖其通信结果才能进行,因此无法通过通信计算重叠来消除其影响。鉴于此,后续优化技术将聚焦于TP AllReduce的优化效果。

### 3.2 拓扑感知的集合通信调度

集合通信算法的目标是完成GPU之间的数据传输,并最大化服务器内和服务器间带宽的利用率。现有集合通信库引入了一种基于环形通信路径的AllReduce算法<sup>[4]</sup>,该算法可以充分利用服务器内和服务器间的带宽。具体过程包括3个步骤:

步骤1:构建一个服务器内的图 $G(N,E)$ 。其中,节点集 $N$ 包含服务器内的所有GPU和NIC,链路集 $E$ 包括GPU-GPU和GPU-NIC链路,每条链路 $e_{ij} \in E$ 均有一个数值表示节点 $i$ 和 $j$ 之间的可用带宽。例如,在图4中,NIC0与GPU0互联,NIC0和GPU0之间的可用带宽为 $\min(bw_{PCIe}, bw_{NIC})$ ,其中 $bw_{PCIe}$ 和 $bw_{NIC}$ 是外围组件快速互连(PCIe)链路和NIC的带宽。此外,由于所有GPU都通过NVSwitch互联,因此每对GPU之间的双向带宽都是NVLink带宽。

步骤2:基于 $G(N,E)$ 搜索连接所有GPU的通信路径,并确定用于连接其他服务器的NIC。每条通信路径均可以表示为一个链式列表 $p_{intra}^i$ ,其中 $i \in [1, I]$ 表示第 $i$ 个服务器。每条路径 $p_{intra}^i$ 的头和尾用于连接其他服务器的NIC,可以表示为 $p_{intra}^i \rightarrow head$ 和 $p_{intra}^i \rightarrow tail$ 。每条路径 $p_{intra}^i$ 的其他元素表示服务器内的所有GPU。例如,图4中Server1的可用路径是NIC0→GPU0→GPU1→GPU2→GPU3→NIC3。

步骤3:根据指定的服务器顺序连接在步骤2中获得的每条路径。图4中的橙色箭头表示两个服务器之间的完整路径,即NIC0→GPU0→……→GPU3→NIC3→NIC0→GPU0→……→GPU3→NIC3。需要说明的是,图4中的通信路径仅占Server1中NIC0和NIC3的一个方向带宽。

基于环形通信路径的AllReduce算法已被证明是最优的,适用于LLM训练中的大带宽集合通信<sup>[4-5]</sup>。为了确保基于环算法的性能,服务器之间的通信需要限制在同一轨道内,以避免较长的通信长度和不均匀的负载。具体来说,跨轨道流量需要通过骨干交换机和其他轨道上的Leaf交换机。此外,跨轨道流量可能与现有流量发生冲突,导致两者速度减慢,



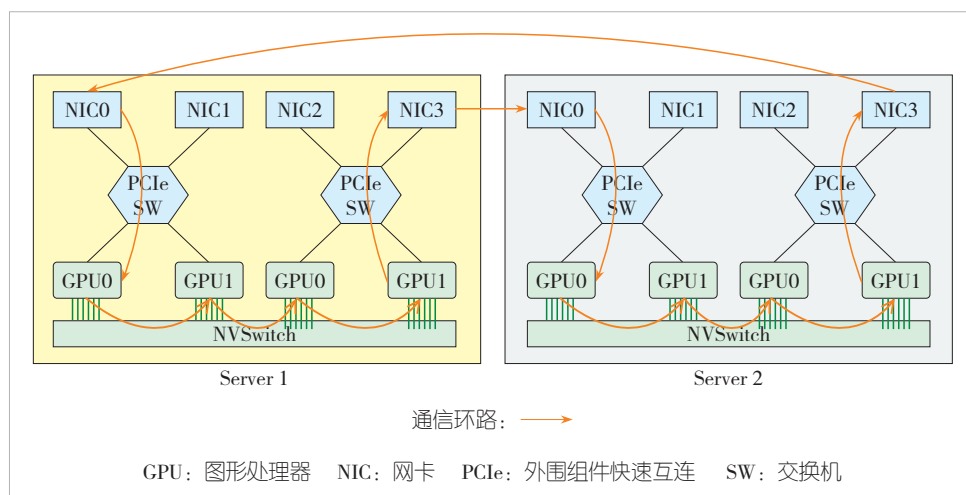


图4 服务器间构建通信环路

从而导致集合通信性能下降。然而, NCCL在通信路径规划的具体实现中并没有考虑轨道网络架构的影响。例如, NCCL按照服务器顺序构建了一个通信路径: Server1 (NIC0-GPU0-GPU7-NIC7) → Server2 (NIC0-GPU0-GPU7-NIC7) → Server3 (NIC0-GPU0-GPU7-NIC7) → Server4 (NIC0-GPU0-GPU7-NIC7)。其中, 节点间连接都是通过跨轨道的NIC0 (轨道0) 和NIC7 (轨道7) 互联实现的, 这会导致所有跨节点流量均为跨轨道流量。

由于交换机的端口数量有限 (在我们的测试场景中, 一个交换机有128个端口), 每个Block只能容纳32个配备有8个NIC的服务器。随着集群规模的增加, Block的数量不可避免地会增加。Block之间通过Spine交换机实现互联, 因此跨Block的流量必须通过Spine交换机。跨Block流量增多会导致服务器之间的通信时延增加、Spine层的交换机拥塞概率增大。因此, 在决定服务器之间的通信顺序时, 需要考虑服务器在Block中的分布。然而, NCCL无法感知服务器在网络架构中的分布情况, 因此通常默认根据给定的服务器列表顺序确定服务器通信顺序。然而, 这会额外产生大量不必要的跨Block流量。例如, 如果遵循给定的服务器顺序, Server1 (Block1) → Server3 (Block2) → Server2 (Block1) → Server4 (Block2) → Server1 (Block1), 跨Block流量的比例为100%。然而, 如果同一Block中的服务器完成顺序连接, 跨Block流量的比例将下降到50%。

为了解决上述集合通信算法的问题, 本文提出了一种分层拓扑感知通信路径规划算法, 以优化基于环形通信路径的AllReduce算法, 使其适应多轨道网络架构并减少跨轨道流量。具体来说, 首先确保所有服务器的最佳通信顺序, 然后使用同一轨道中的NIC来连接服务器内的通信环, 以避免跨

轨道流量。基于此, 我们首先找到所有服务器所在的Block, 并将属于第 $k$ 个Block的所有服务器聚类到集合 $block_k$ 中; 然后, 对集合 $\{block_k, k \in \{1, \dots, K\}\}$ 中的所有服务器进行排序, 以获得最佳服务器通信序列, 其中 $K$ 是集群中的块数; 之后, 遍历所有轨道 $\{rail_j, j = \{1, \dots, J\}\}$ , 在每个服务器上找到属于同一轨道 $rail_j$ 的NIC, 并将其作为服务器内通信路径的头和尾, 以避免跨轨道流量; 最后, 将所有服务器内通信路径 $\{p_{intra}^i, i \in \{1, \dots, |\Psi|\}\}$ 与 $p_{intra}^i - 1 \rightarrow next = p_{intra}^i \rightarrow head$ 连接, 形成一个完整的通信环, 其中 $|\Psi|$ 是最佳服务器通信序列 $\Psi$ 的长度。

### 3.3 RoCE 和 NVLink 并行通信优化

LLM模型通过TP和PP拆分为多个部分, 并部署在不同的GPU上。每个TP组和PP组利用集合通信 (例如AllReduce) 在模型层内和层间实现参数传输。由于每个模型层在前向和后向过程中需要4次TP AllReduce, 根据GPT-3175B的并行设置, TP AllReduce的总次数可以达到6 720次, 其中每次TP AllReduce的通信量可以达到224 MB。为了减少通信延迟, 现有工作将每个TP组限制在服务器内, 以充分利用高速互联 (例如NVLink)。然而, 如图3所示, 我们的GPU服务器不仅配备了高速互联, 还为服务器中的每个GPU配备了一个NIC。这意味着服务器内的TP AllReduce可以通过NVLink以及绑定到GPU的NIC完成。对此, 我们提出了一种异构通信优化方案, 该方案可利用空闲RoCE网络来增强服务器内的集合通信 (例如TP AllReduce), 进一步提高系统训练性能。

为了实现服务器内和服务器间的多路径并行通信, 常规做法是将发送的消息拆分为多个块, 然后轮询每条路径发送块, 并在路径末端组合块以获得最终结果。然而, TP AllReduce需要在每个GPU上同步一次传输的结果以进行加法操作。整个集合通信的延迟取决于最慢的路径。需要注意的是, NVLink和RoCE网络的传输延迟分别为200 ns和4  $\mu$ s。这导致通信性能受限于RoCE网络的延迟。这一问题的根本缘由在于, 两条通信路径没有完全解耦, 需要在每次传输后完成同步操作。因此, 我们提出在集合通信粒度上切片TP

AllReduce, 将每个GPU上要传输的消息分为两个部分, 分别用于NVLink通道和网络通道。每部分消息可以独立通过NVLink或网络通道传输, 以获得最终的集合通信结果。图5展示了在两个通道上并行传输数据的操作原理。GPU0中的蓝色方块数据可以通过NIC0和NIC1之间的RoCE网络传输到GPU1, 并与GPU1中的相应橙色方块完成加法操作。GPU2将重复相同的操作以获得最终AllReduce结果。带有蓝色条纹的方块也可以实现NVLink通道上的相同AllReduce过程。这两个通信过程完全独立, 以确保NVLink和网络通道都能实现线速传输。

此外, 还有一个重要的问题是可以将多少数据卸载到网络通道上, 以便提高服务器内集合通信的性能。因此, 卸载到网络通道的数据, 需要在NVLink通道完成剩余数据传输任务之前完成传输。显然, 基于网络通道和NVLink通道的静态延迟比值来确定卸载到网络通道的数据量是不可能的。这是因为背景流量和异步通信等因素导致两个通道的状态动态变化。为此, 我们设计了一种动态滑动窗口机制, 以便动态调整在两个通道上传输的数据量。具体来说, 我们首先将发送的消息分为 $m$ 个块, 并根据NVLink通道和网络通道的带宽计算传输窗口的大小(即 $Win_{NVLink}$ 和 $Win_{Net}$ ), 以分配每个通道上的数据。例如, 如果NVLink通道和网络通道的AllReduce带宽分别为100 GB/s和50 GB/s, 那么 $Win_{NVLink}$ 和 $Win_{Net}$ 可分别设置为2个块和1个块。如图5所示, NVLink通道和网络通道的初始传输窗口分别被设置为待传输数据的头和尾。之后, 两个通道开始并行传输相应窗口中的数据。当

任一通道的窗口数据传输完毕时, 窗口滑动 $Win_{NVLink}$ 或 $Win_{Net}$ 以准备下一次传输。持续重复这一过程, 直到两个通道传输的块数累计达到 $m$ 个块。

### 3.4 网络流量规划

网络流量规划的目标是, 通过均衡交换机间的负载分配避免潜在的瓶颈或拥塞。为实现高可用性, NIC采用了双端口上联两个Leaf交换机的模式。为了充分利用NIC和两个Leaf交换机之间的带宽, 我们计划在每个源地址与目的地址间建立至少两条流(例如设计2个QP, 实际使用4个QP)。此外, 交换机通常采用ECMP的方式进行流级负载均衡, 并根据五元组哈希值将流量分配到不同路径。实际上, 大多数拥塞问题源于多个流被哈希到同一路径上。

#### 1) 路由注入策略的选择

一种可能的方法是通过向交换机注入路由策略来规划流路径。基于策略的路由(PBR)可通过定义五元组流的下一跳建立访问控制表(ACL)策略, 并通过全局规划最小化流路径冲突。然而, 该方法存在两个缺陷: 一是, 交换机的ACL规则数量通常有限, 而全局规划需在每个交换机部署大量的ACL规则; 二是, 管理过多ACL规则对网络维护人员而言是巨大负担。因此, 我们通过特定流源端口来规划路径以最小化冲突。

#### 2) 源端口号的确定

若未理解ECMP路由算法, 通信库只能借助随机端口号使流随机选择路径。这种方法无法保证多流均匀分布, 也无法充分利用大规模网络中的可用

路径。实验表明, 此类不均衡可能导致可用带宽下降25%。现有工作<sup>[6]</sup>证明, 基于ECMP哈希函数的线性特性可避免同设备对之间多个QP的路径冲突。假设ECMP哈希函数为 $H(p)$  ( $p$ 为流源端口), 其线性特性表现为 $H(p \oplus \delta) = H(p) \oplus H(\delta) \oplus H(0)$ 。因此, 对于源端口 $p$ 和 $p \oplus \delta$ , 存在如下等式关系:  $H(port) \oplus H(port \oplus \delta) = H(\delta) \oplus H(0)$ 。由此可知,  $H(p)$ 与 $H(p \oplus \delta)$ 是否相等仅由 $\delta$ 决定。为确保 $H(p) \neq H(p \oplus \delta)$  (即两流使用不同路径), 只需要 $H(\delta) \neq H(0)$ 即可。

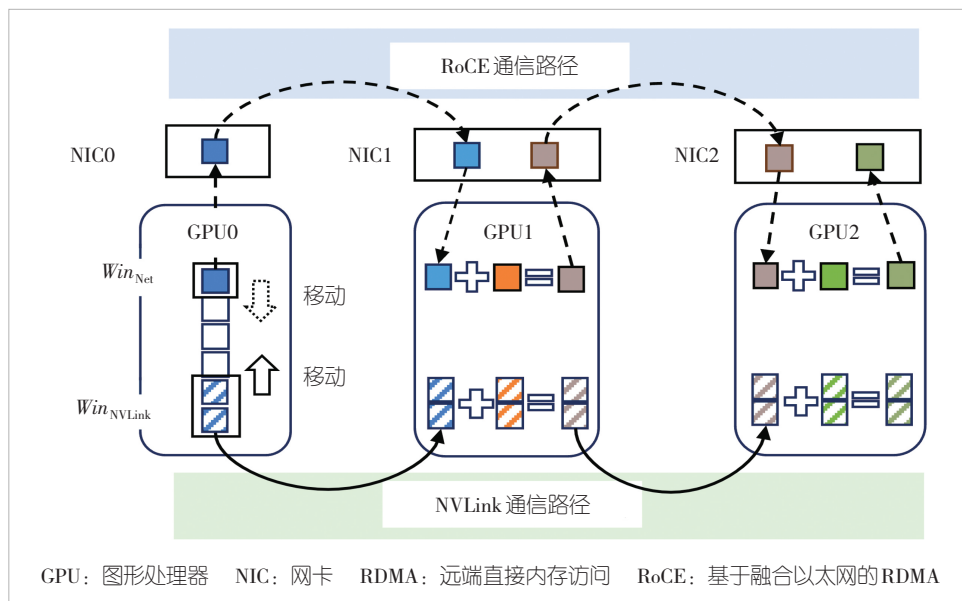


图5 RoCE网络 and NVLink网络并行通信优化

在实际操作中,可通过各交换机的哈希算法模拟器,固定源端口号(如50000)并顺序调整目标端口号,观察哈希值变化。通过遍历目标端口号范围(如1~65536),筛选出4个编译量 $\delta$ (如 $\delta=1\sim 4$ ),使得 $H(\delta)\neq H(0)$ 。这4个目的端口号将分配给4个QP(即4条流),以避免流量冲突。

## 4 动态网络流量调度

尽管路由规划借助源端口多路径分发机制确保了流量的均匀分布,但在实际应用场景中仍存在问题:1)考虑到高性能网络的规模,链路中断导致网络拓扑结构改变的情况普遍存在;2)预规划旨在通过哈希规则将流数量均匀分配到各网络链路,但由于流大小存在差异,当大流量被分配到同一条链路时仍可能会引发拥塞;3)在需要多任务并行部署的云环境中,新任务的到达具有不可预知性。因此,仅依赖流量预规划无法有效解决上述问题。

星脉网络能够通过集中式控制器来应对动态流量调度中的双重挑战:

- 实时拥塞检测:基于分布式探针和流量特征分析,识别毫秒级拥塞事件;
- 全局动态路由规划:利用在线优化算法动态调整流量分配策略。

星脉网络的核心创新在于将调度周期压缩到一个LLM训练迭代周期(典型值为10~30 s)内,在当前迭代内完成下一个迭代的流量调度和资源预留,从而抑制流量拥塞跨训练迭代传递。

### 4.1 网络拥塞监测

基于对LLM训练过程日志的分析,我们发现网络事件可以分为两类:78%表现为连接中断,22%表现为因链路拥塞导致的性能下降。实时监控网络状态对于及时检测拥塞和定位受影响数据流具有关键作用。

在高性能网络场景中,LLM训练流量呈现高突发性,链路拥塞持续时长通常在数十毫秒至数百毫秒量级。实验证实,秒级粒度的链路状态观测无法有效捕捉实际发生的拥塞事件(这些拥塞已显著影响训练收敛性)。然而,采用更细粒度的统计窗口(如亚秒级采样)会引发两个技术难题:一是,系统日志量存呈指数级增长;二是,现有硬件设备不支持高频数据采集。因此,本文提出基于ECN的拥塞感知方法,通过优化拥塞指标敏感性解决上述矛盾。

ECN是数据中心传输控制协议、数据中心量化拥塞通知协议等拥塞控制协议的核心机制<sup>[7]</sup>。在实际部署中,系统以10 s为周期统计交换机端口的ECN标记数量,并定义三级告

警规则:若1 min内ECN标记数连续3次超过500阈值,则触发中级告警;若相邻LLM训练阶段间隔(典型值为10~30 s)内累计告警次数 $\geq 5$ 次,则触发高级告警。该设计在保证采样频率与LLM训练流量周期对齐的同时,减小了误告警出现的概率。

### 4.2 动态场景下全局路由规划

在检测到由流量冲突引起的中、高级ECN告警后,我们通过sFlow流量采样技术<sup>[9]</sup>和Telemetry网络遥测技术<sup>[10]</sup>识别待重路由的流量,并通过哈希注入和路径模拟来选择最优路径。具体流程如下:

1) 流量识别:当交换机触发ECN告警时,sFlow从告警端口采集所有流量的五元组信息,并以10 s为一个周期持续执行6次连续状态查询。按平均带宽降序排序后,选取在6个周期中出现 $\geq 2$ 次的Top  $N$ 流量,记为集合 $F$ 。

2) 流量上下文解析:对于每个流 $Flow_i \in F$ ,提取其源GPU $_i$ 、连接的Leaf交换机 $leaf_k$ 、在 $leaf_k$ 上的出端口 $GE_{ki}$ ,并构建集合 $C$ ——包含 $leaf_k$ 上所有与 $Flow_i$ 具有相同目标IP但源IP不同的流(不包含 $Flow_i$ 自身)。由于交换机上的流量路由控制粒度不是针对特定流的,而是针对具有相同目标IP的所有流量,因此,我们需要考虑集合 $C$ 中所有流量在新路由策略下是否可以避免流量冲突。

3) 下一跳探测:通过遥测技术查询 $leaf_k$ 上 $Flow_i$ 目标IP可用的下一跳(排除 $GE_{ki}$ )及其空闲带宽,并按空闲带宽降序排列形成有序列表 $L_o$ 。

4) 路径重计算:

• 哈希模拟:提取交换机的五元组哈希算法,构建路径计算模拟器。

• 路径映射:对于集合 $C$ 中的每个流 $Flow_k$ ,使用模拟器计算其在下一跳集合中命中的出口端口,并确定该端口连接的 $Spine_j$ 交换机。

• 递归验证:基于 $Flow_k$ 目标IP查询 $Spine_j$ 的路由表,确定其下一跳集合 $H_j$ ,随后递归式地找出 $Flow_k$ 经过的所有物理链路路径。

• 冲突检测:在确定 $Flow_k$ 的所有物理链路路径后,执行链路状态检查,具体包括:(1)优先从控制器内存中获取实时链路占用数据;(2)若内存无数据,则通过Telemetry系统实时查询;(3)若 $Flow_k$ 带宽叠加后,链路的总占用率超过其容量的75%,则终止当前迭代并返回步骤。

5) 路由更新:将验证通过的路由策略同步交换机控制面,完成ECMP表项更新。



实验数据表明,该机制可显著缩短网络拥塞持续时间。在实际测试中,拥塞时长降幅超过80% ( $p < 0.01$ ),满足AI训练场景对微秒级拥塞响应的需求。

## 5 星脉网络性能评估

### 5.1 评估环境设置

我们通过一个NVIDIA GPU集群对星脉网络进行性能评估。该集群的网络采用8轨道网络架构设计。每块GPU的NVLink高速互联总线通过NVSwitch交换芯片与机内其他7个GPU互联,并同时与一个双端口Mellanox ConnectX-6/7智能网卡绑定。跨机网络基于RoCEv2协议构建,通过PFC流控机制实现无损传输。负载均衡采用GOR策略,同时引入DCQCN,并结合动态ECN配置<sup>[11]</sup>,实现网络拥塞控制的管理。

### 5.2 TCCL的性能评估

我们采用标准NCCL-TEST测试工具<sup>[5]</sup>对TCCL和开源NCCL(评估中采用2.17.1版本)进行性能评估。实验采用

NCCL-TEST定义的全规约总线带宽(简称Busbw)作为核心指标,其计算公式为:  $\text{Busbw} = \frac{2(n-1)}{n} \times \text{Algbw}$ ,其中, $n$ 表示GPU的数量,Algbw为算法带宽(可通过通信信息与集合通信耗时)。总线带宽Busbw能更准确地反映集合通信对物理带宽的实际利用率,因此,后续实验结果均采用该指标进行量化分析。

图6(a)中蓝色线的性能上限表示,在1 Gbit/s通信量下AllReduce的集合通信带宽的理论上限(单机8个网卡带宽之和,即200 Gbit/s)。在相同条件下,TCCL几乎达到该理论极限,较NCCL提升了22%。对于64 MB及以上的通信量,采用异构通信优化的TCCL始终优于NCCL,性能提升幅度为9%~25%。然而,当通信量小于64 MB时,TCCL与NCCL性能相近。这是因为此时性能受限于通信网络时延,所有的数据都会优先通过NVLink网络传输(NVLink的传输时延为200 ns,远小于RoCE的传输时延2  $\mu\text{s}$ )。这意味着TCCL方案效果和NCCL方案一致,因此两者最终的AllReduce Busbw性能也一致。异构并行策略优化不仅对AllRe-

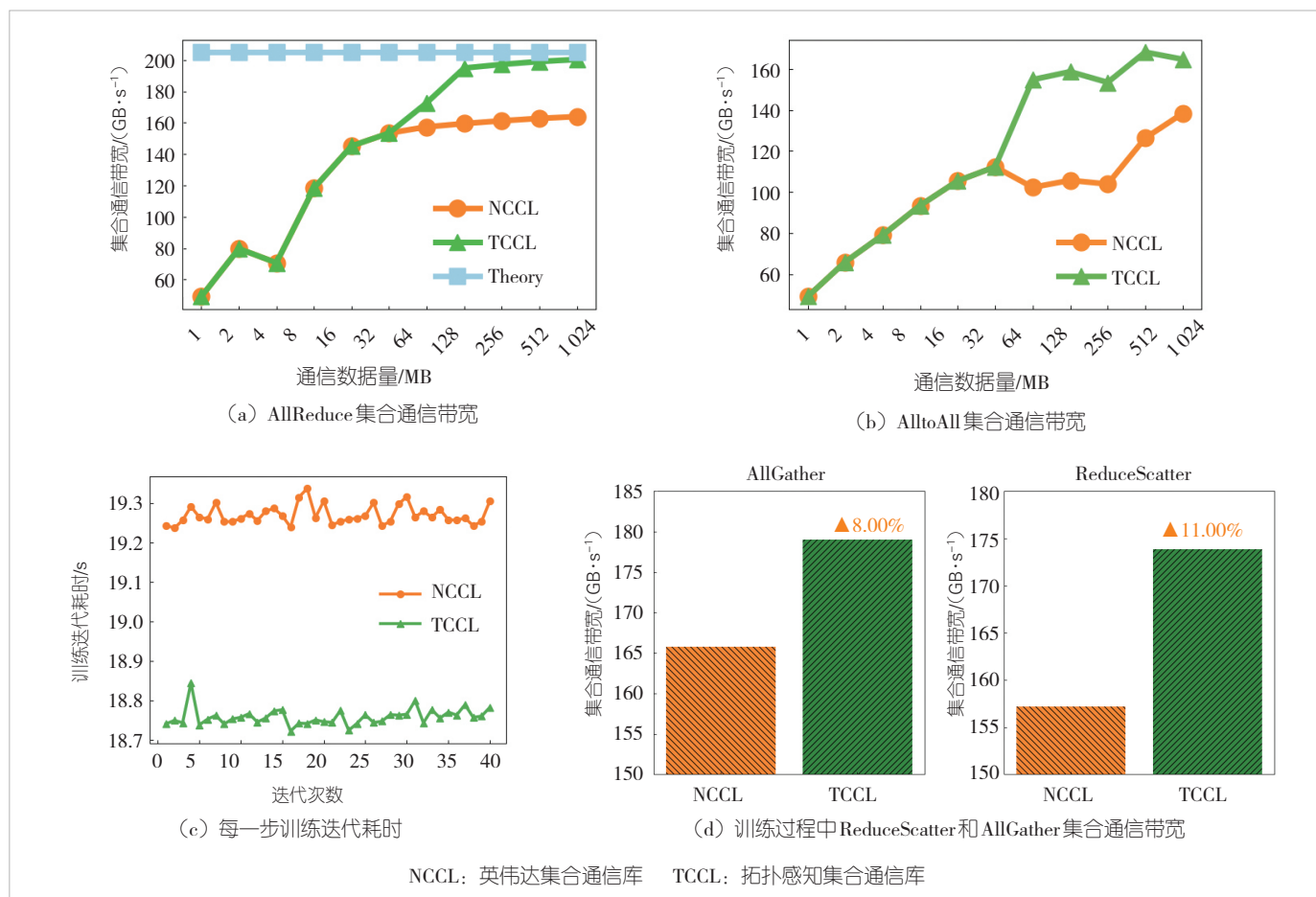


图6 TCCL异构并行通信优化的性能评估

duce 场景有效,还能提升其他集合通信操作(如 AlltoAll)的性能。如图 6 (b) 所示,在 1 GB 通信量下采用异构并行优化的 AlltoAll 性能提升了 20%。

为进一步验证该优化的有效性,我们在 64 台 H800 服务器上开展对比实验,分别使用 TCCL 与 NCCL 对类 GPT 模型进行训练,并对两者的耗时情况进行对比分析。如图 6 (c) 所示,采用异构通信优化的 TCCL 在每轮训练迭代中平均节省约 2.5% 的时间。此外,运行时对 TP AllReduce 的带宽监测显示,AllGather 和 ReduceScatter 操作分别实现了 8% 和 11% 的性能提升。这些结果充分证明了异构通信优化在不同类型集合通信操作中的广泛有效性。

图 7 展示了拓扑感知路由在不同节点规模下的性能增益(测试通信量大小为 1 GB)。由图 7 (a) 可知,随着节点数量增加,拓扑感知路由带来的性能提升显著增长:当节点规模达到 140 时,性能提升幅度可达 13%。这是由于节点规模扩大导致通信节点数量增加,NCCL 方案会产生跨 Spine 交换流量,加剧流量负载不均衡。因此,在大规模集群中,需通过精细的流量规划来避免非必要的跨 Spine 交换机通信。

随着节点数量的增加,拓扑感知路由的性能增益也在增加。当节点数量达到 140 时,系统性能将提升 13%。这是因为,随着节点数量的增加,涉及的块数量也随之增加。NCCL 引入了跨 Spine 流量,导致流量负载更加不平衡。因此,在大规模集群中,需要仔细规划机器之间的流量,以避免不必要的跨 Spine 流量。

为了评估分层拓扑亲和流量规划对集体通信稳定性的影响,我们从 100 台机器中随机选择 4 台进行 200 次重复实验,对比 TCCL 和 NCCL 的 AllReduce Busbw 性能(见图 7 (b))。统计分析表明,TCCL 与 NCCL 的平均集合通信性能均稳定在

11.5 GB/s。NCCL 在 200 次实验中的标准差为 1.26,显著高于 TCCL 的 0.31。该差异源于 NCCL 存在大量的跨 Spine 流量,这种情况增加了负载不均衡概率,从而导致通信性能的波动。

在 GPT3 模型训练任务中部署 TCCL 后,我们持续监测一周的跨 Spine 流量(结果见图 7 (c))。相比于未采用 TCCL 的方案,该方案的跨 Spine 流量减少了 75%。流量优化有效保障了网络负载均衡,同时使 GPT3 的样本训练吞吐量从 NCCL 方案的 45 个/s 提升至 49 个/s。

### 5.3 GOR 的性能评估

我们首先评估了 GOR 控制器在 64 MB ~ 1 GB 通信数据量下对 AllReduce 任务的动态流量调度能力。实验表明,无论通信量多大,都会触发流量拥塞告警。这是因为,集合通信的流数量由 GPU 数量和 QP 队列对数决定。同时我们发现,大部分拥塞告警是由哈希冲突引发的。为此,我们启用 GOR 对所有冲突流量进行调度,直至消除所有拥塞告警。图 8 (a) 展示了调度前后的测试结果,可以看出 AllReduce 的 Busbw 性能提升了 8.4% ~ 18.8%。

我们随后在 AlltoAll 场景中评估了 GOR 控制器的性能。由于 AlltoAll 涉及所有参与网卡间的通信并产生大量数据流,其拥塞概率会显著增加,同时流量调度也更为复杂。图 8 (b) 显示,当通信量为 64 MB 且关闭 GOR 时,AlltoAll 的 Busbw 达到最低值。这是因为该配置触发了 DCQCN 水位线机制并产生大量 ECN,导致网卡降速。在启用 GOR 后,控制器将拥塞流量调度至相对空闲的路径,使 Busbw 性能提升 77.4%,如图 8 (b) 所示。

在运营网络验证中,图 8 (c) 显示启用 GOR 调度后,集群拥塞告警数量与持续时间均下降 80%。系统稳定后,每

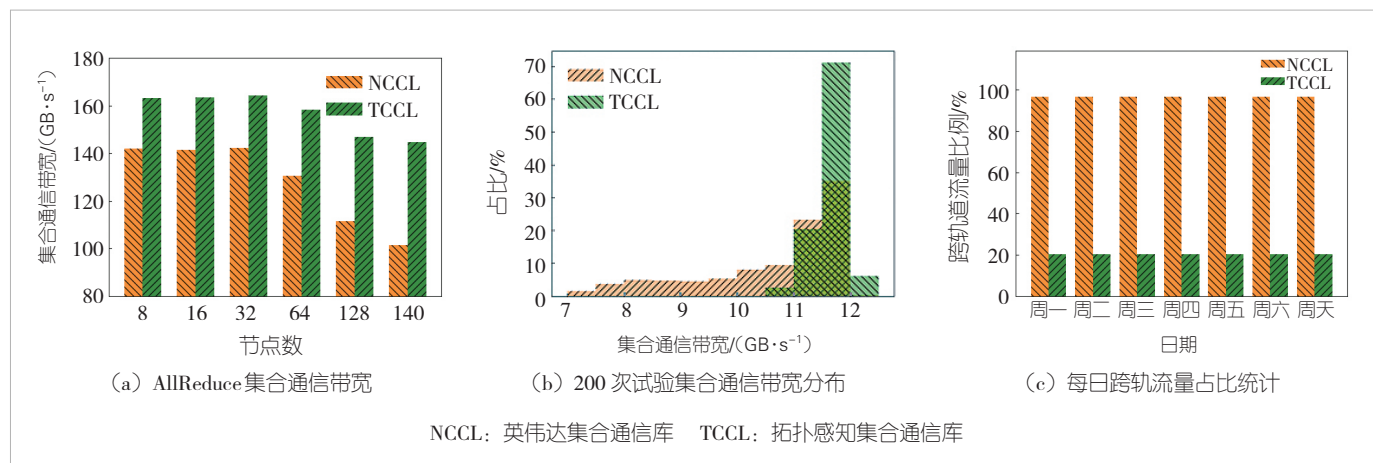


图 7 TCCL 拓扑感知流量路径规划的性能评估

日告警量维持在10次左右,每次可通过1~2次调度在10 s内解决。告警数量的波动源于新任务到达时的流量模式变化,而持续观测结果验证了动态流量调度的稳定性。

图9(a)展示了单端口场景下GOR的拥塞消除效果。调度后交换机出口的ECN计数器归零,表明拥塞即时解除。图9(b)则呈现了多并发流导致的严重拥塞处理过程。通过三轮调度(将最大流量依次迁移至Link1和Link2),ECN通知量从初始的10 000以上逐步降至500以下,最终有效消除了链路拥塞。

6 结论

本文介绍了星脉网络,一个用于万卡GPU集群互联的高性能无损网络实践案例。星脉网络采用端侧集合通信库与全局路由控制联合优化方案,使AllReduce和AlltoAll通信带宽性能相较于公开最优方案(NCCL)分别有25%和22%的提升。网络测试结果表明,星脉网络能够有效管理LLM大模型训练中的大带宽和高突发的流量,把网络拥塞概率降低到1%以内,使网络拥塞时长下降80%。

参考文献

[1] YENDURI G, RAMALINGAM M, CHEMMALAR S G, et al. Generative pre-trained transformer: a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions [EB/OL]. [2025-03-10]. <https://arxiv.org/abs/2305.10435v2>

[2] GAO Y, LI Q, TANG L, et al. When cloud storage meets RDMA [EB/OL]. [2025-03-10]. <https://www.usenix.org/system/files/nsdi21-gao.pdf>

[3] SONG J, YIM J, JUNG J, et al. Optimus-CC: efficient large NLP model training with 3D parallelism aware communication compression [EB/OL]. [2025-03-10]. <https://arxiv.org/abs/2301.09830v1>

[4] NVIDIA Developer. NVIDIA collective communications library (NCCL) [EB/OL]. [2025-03-10]. <https://developer.nvidia.com/nccl>

[5] Github. NCCL tests [EB/OL]. [2025-03-10]. <https://github.com/NVIDIA/nccl-tests/tree/master>

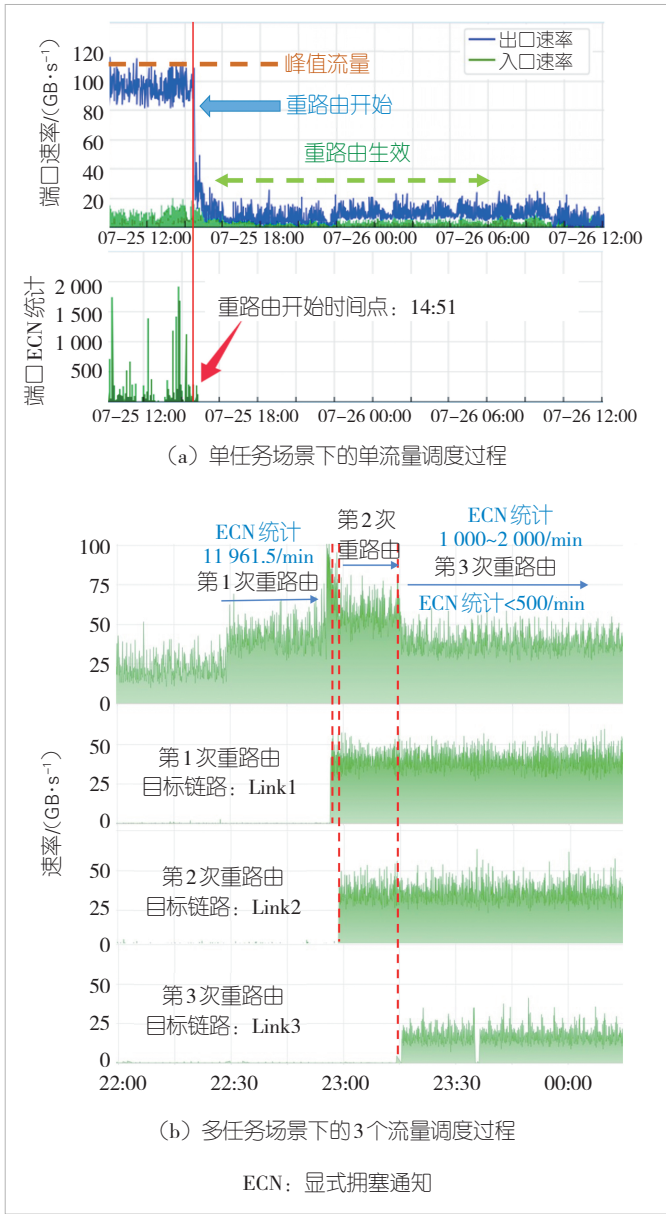


图9 全局优化路由器(GOR)消除流量拥塞过程

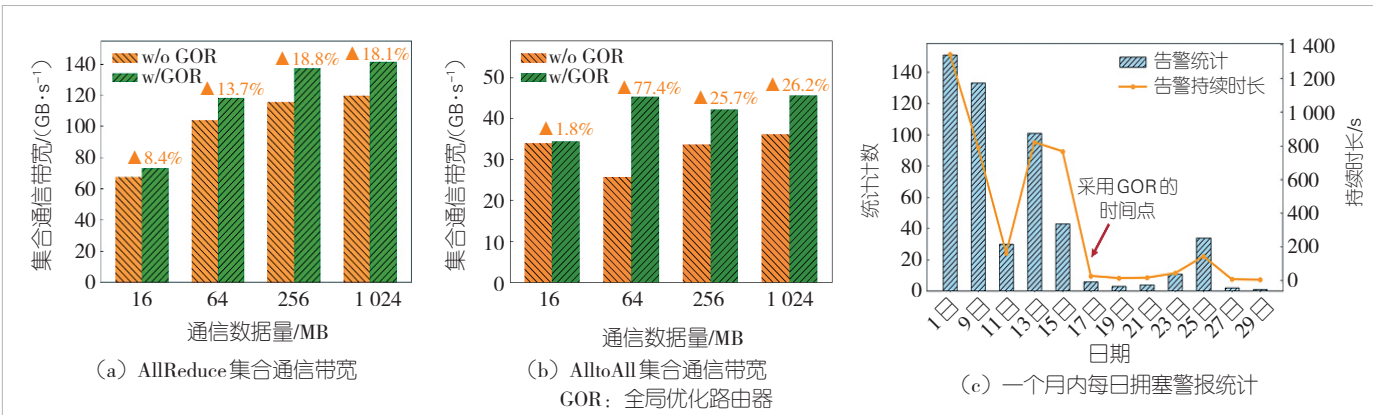


图8 使用GOR控制器的动态流量调度性能评估

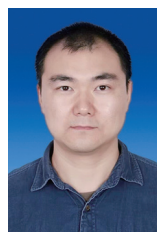


- [6] ZHANG Z, ZHENG H, HU J, et al. Hashing linearity enables relative path control in data centers [EB/OL]. [2025-03-10]. <https://www.usenix.org/conference/atc21/presentation/zhang-zhehui>
- [7] KUMAR G, DUKKIPATI N, JANG K, et al. Swift: delay is simple and effective for congestion control in the datacenter [C]//Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication. ACM, 2020: 514–528. DOI: 10.1145/3387514.3406591
- [8] NARAYANAN D, SHOEYBI M, CASPER J, et al. Efficient large-scale language model training on GPU clusters using megatron-LM [C]//Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. ACM, 2021: 1–15. DOI: 10.1145/3458817.3476209
- [9] WANG M, LI B, LI Z. sFlow: towards resource-efficient and agile service federation in service overlay networks [C]//Proceedings of 24th International Conference on Distributed Computing Systems, IEEE, 2004: 628–635. DOI: 10.1109/ICDCS.2004.1281630
- [10] YU M L. Network telemetry [J]. ACM SIGCOMM computer communication review, 2019, 49(1): 11–17. DOI: 10.1145/3314212.3314215
- [11] YAN S Y, WANG X L, ZHENG X L, et al. ACC: automatic ECN tuning for high-speed datacenter networks [C]//Proceedings of the 2021 ACM SIGCOMM 2021 Conference. ACM, 2021: 384–397. DOI: 10.1145/3452296.3472927

## 作者简介



**李宝嘉**, 腾讯数据中心网络架构师; 主要负责AI大模型超算网络架构方案论证和设计、训练架构加速方案优化以及集合通信优化。



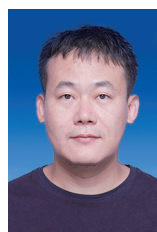
**何春志**, 腾讯数据中心网络架构师; 主要负责腾讯高性能计算网络的架构设计、集合通信库/网络协议研究, 以及大模型训练推理业务与高性能网络的联合协同优化。



**夏寅贇**, 腾讯网络首席架构师; 主导腾讯星脉网络系统的设计研发工作, 构建从自研软硬件系统到端到端AI集群运维的高性能网络系统, 支撑腾讯多个万卡AI集群的快速建设与高效运行。



**何泽坤**, 腾讯基础网络中心总监, 并担任开放数据中心委员会(ODCC)网络工作组组长; 长期深耕数据中心网络、骨干网络架构, 主导腾讯全球数据中心互联网络及与运营商互联网络的设计, 近年来聚焦AI算力网络协同创新, 牵头开放数据中心委员会ETH-X开放超节点与MegaScaleOut项目。



**王晓亮**, 南京大学计算机学院副教授; 长期从事网络体系结构的研究工作; 曾获APNET、BIGCOM、EuroSys最佳论文奖, 并获得2019年江苏省科技进步奖一等奖、2023年江苏省科技进步奖二等奖; 发表论文50篇。

# 智算中心 Scale-Out 网络的演进及 GSE 的实践



## Evolution of Scale-Out Network in Intelligent Computing Centers and Practice of GSE

程伟强/CHENG Weiqiang<sup>1,2</sup>, 李新双/LI Xinshuang<sup>3</sup>,  
白艳/BAI Yan<sup>2</sup>, 吕勇/LIU Yong<sup>3</sup>

(1. 东南大学, 中国 南京 211189;

2. 中国移动通信有限公司研究院, 中国 北京 100053;

3. 中兴通讯股份有限公司, 中国 深圳 518057)

(1. Southeast University, Nanjing 211189, China;

2. The Research Institute of China Mobile, Beijing 100053, China;

3. ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.202502003

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250425.1536.003.html>

网络出版日期: 2025-04-27

收稿日期: 2025-03-05

**摘要:** 探讨了人工智能 (AI) 大模型时代智算中心网络面临的技术挑战, 重点分析了传统互联网协议 (IP) 网络在负载均衡和突发流量处理方面的局限性, 并对基于以太网融合远程直接内存访问 (RoCE) 的优化与网络架构重构两种技术路线。研究聚焦中国自主研发的全调度以太网 (GSE) 技术, 详细阐述了其核心技术: 基于报文容器 (PKTC) 的负载均衡机制和动态全调度队列 (DGSQ) 端到端拥塞控制技术。这些技术有效解决了智算网络中的流量极化和拥塞丢包问题。同时, 系统分析了 GSE 网络设备在接口设计、转发引擎和队列管理等关键环节的创新架构, 论证了 GSE 技术在构建高带宽、低时延、无阻塞新型网络方面的技术优势, 为智算中心网络演进提供了重要参考。

**关键词:** AI 大模型; 智算中心; Scale-Out; GSE; RoCE; 负载均衡; 拥塞避免

**Abstract:** The technical challenges faced by intelligent computing center networks in the era of large-scale artificial intelligence (AI) models are discussed, focusing on analyzing the limitations of traditional Internet Protocol (IP) networks in load balancing and burst traffic handling. Two technical approaches are compared: optimization based on remote direct memory access over converged Ethernet (RoCE) and network architecture reconstruction. The research centers on China's independently developed global scheduling Ethernet (GSE) technology, detailing its core innovations: the packet container (PKTC)-based load balancing mechanism and the dynamic global scheduling queue (DGSQ) end-to-end congestion control technology, which effectively addresses traffic polarization and congestion packet loss in intelligent computing networks. Additionally, it systematically analyzes the innovative architecture of GSE network equipment in key areas such as interface design, forwarding engines, and queue management, demonstrating the technical advantages of this approach in building high-bandwidth, low-latency, and non-blocking next-generation networks. The findings provide important insights for the evolution of intelligent computing center networks.

**Keywords:** AI large-scale model; intelligent computing center; Scale-Out; GSE; RoCE; load balance; congestion avoidance

**引用格式:** 程伟强, 李新双, 白艳, 等. 智算中心 Scale-Out 网络的演进及 GSE 的实践 [J]. 中兴通讯技术, 2025, 31(2): 14-20. DOI: 10.12142/ZTETJ.202502003

**Citation:** CHENG W Q, LI X S, BAI Y, et al. Evolution of Scale-Out network in intelligent computing centers and practice of GSE [J]. ZTE technology journal, 2025, 31(2): 14-20. DOI: 10.12142/ZTETJ.202502003

### 1 AI 大模型对网络的挑战

以生成式预训练变换器 3.0 (GPT 3.0) 为代表的大模型展现惊人的能力后, 人工智能 (AI) 呈现向海量参数大模型方向发展的技术趋势。随着算力的提升与数据资源的

不断扩充, 大模型的参数量级将继续扩大。目前, 已有多个大模型的参数规模超过了万亿级别。

当前, 单独的芯片和存储芯片已无法满足 AI 大模型对参数量和计算量的需求, 形成了制约 AI 技术发展的“算力墙”和“存储墙”两大瓶颈。为此, 业界普遍采用通过多计算节点构建高性能集群的方案, 以整合分布式计算能

基金项目: 国家重点研发计划项目 (2024YFB2906600)

力和存储资源,从而突破算力和存储的双重限制。这种基于集群的分布式架构已成为应对上述挑战的主流解决方案<sup>[1]</sup>。

智算集群中节点间网络的通信效率直接影响集群的整体吞吐量和性能。AI大模型训练业务的网络流量具有以下特征:流数量少(低熵)、单流带宽高(大象流)、同步突发(Incast)等,这对传统基于以太网的IP网络架构提出了两大挑战<sup>[2-3]</sup>:

挑战1:传统基于流的等价多路径路由(ECMP)负载均衡技术在流数量较少时存在局限性,会导致交换网络中出现流量极化现象,从而造成链路负载不均。具体表现为部分链路拥塞而其他链路利用率不足,这会降低整体网络吞吐量,如图1所示。

挑战2:在集群节点通信过程中,当源端在不了解目的端接收能力的情况下持续发送数据,会形成分布式训练中典型的多对一通信模式。这种模式产生的大量Incast流量将导致网络设备队列缓存出现瞬时突发,进而引发拥塞甚至丢包问题,最终造成应用时延增加和吞吐量下降,如图2所示。

因此,如何构建适配大模型算力的高性能网络,突破现

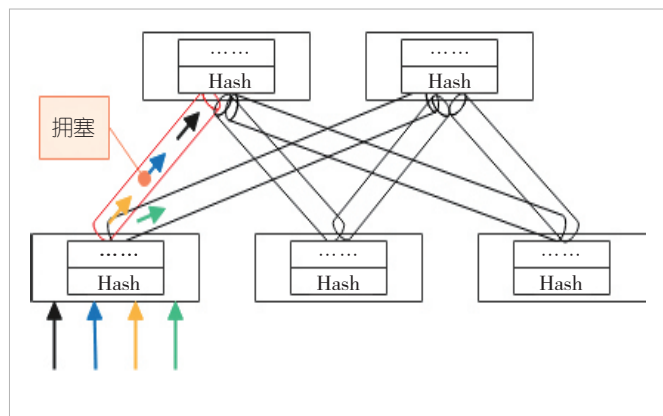


图1 等价多路径路由选路不均造成网络拥塞

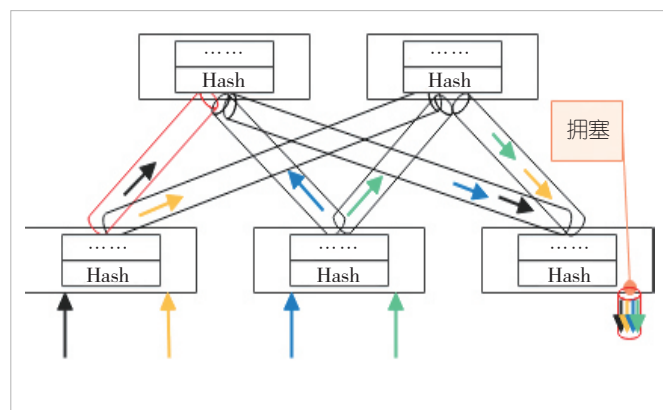


图2 Incast流量造成网络拥塞

有Scale-Out架构的瓶颈,已成为智算时代亟待解决的关键技术难题。

## 2 Scale-Out网络的优化探索

为应对当前网络挑战,人们在探索过程中提出了多种优化思路。根据底层转发优化处理方式的不同,这些方案可归纳为两条发展路线:优化路线与重构路线,具体如图3所示<sup>[4-5]</sup>。

### 1) 基于RoCE的优化

在传统基于融合以太网的远程直接内存访问(RoCE)技术的基础上进行创新优化,通过引入新技术可以改进智能计算Scale-Out网络架构。该方案保持网络底层转发逻辑不变,使现有网络更好地适配算力流量特征,有效降低Scale-Out网络对计算性能的制约影响。

方式1:纯网络侧的优化。纯网络侧的优化是设备厂商倡导的技术方案。该方案是基于网络设备技术升级的无损网络优化方案,主要包括:(1)通过快速显式拥塞通知(ECN)功能降低队列深度带来的时延影响;(2)采用AI ECN智能调优技术简化复杂的水线参数配置;(3)利用智能全局负载均衡提升少流场景的均衡性能。此类创新技术方案持续涌现,推动着网络性能的不不断提升。

方式2:端网协同的优化。端网协同优化是互联网厂商倡导的技术方案,其核心在于通过终端侧的性能优化及网络状态感知,动态协调计算能力与网络资源,实现系统整体效能的提升。例如:阿里 $\mu$ Fab方案采用智能网络调度机制,通过 $\mu$ Fab-E网卡主动发送探测报文(probe),由 $\mu$ Fab-C交换机动态反馈路径带宽和时延信息,基于这些网络状态数据实现网卡级智能限速及动态路径选择;阿里高精度拥塞控制算法(HPCC)与谷歌CSIG(一种用于网络拥塞控制的协议)方案采用端网协同机制,通过网侧随路采集拥塞状态信息,实现端侧流控参数的精细化调优;腾讯的星脉方案,基于多轨异构亲和和部署策略,结合自研的拓扑感知集合通信库(TCCL),显著提升网络通信性能,已在多个场景成功落地。

### 2) 网络架构重构

RoCE的优化建立在现有网络基础之上,这种方式仅能缓解算力与网络之间的冲突,并不能从根本上解决问题,因此称不上是最优解决方案。为彻底摆脱网络困境,业界各方希望构建全新的网络及底层转发机制,突破无损以太网的性能瓶颈,实现无阻塞、高带宽、超低时延,以契合AI与高性能计算对新型网络的需求<sup>[6]</sup>。

在重构路线方面,有两大主流技术方案在业内具备广泛的影响力:

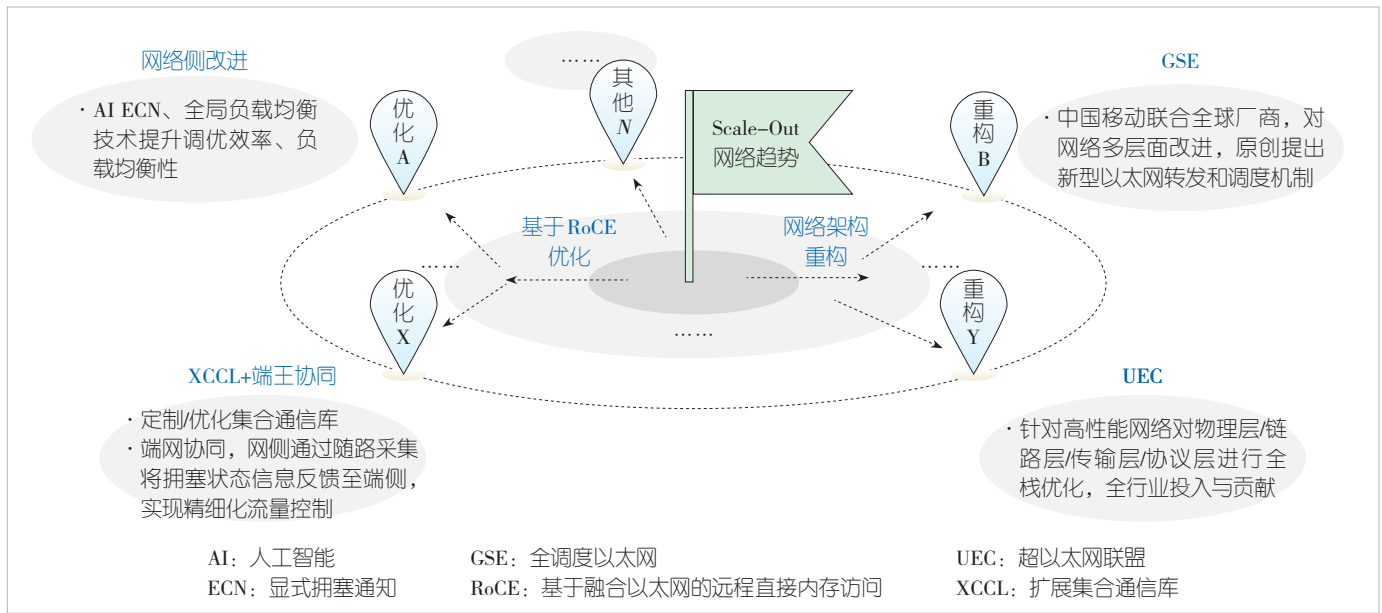


图3 Scale-Out网络路线与实践

(1) 由全球多家颇具影响力的企业主导成立的超以太网联盟（UEC）。该联盟专注于高性能网络，对全通信栈开展深入优化工作，积极整合全行业资源，全力投入并贡献力量。其核心目标在于有效解决大规模数据传输过程中存在的尾部延迟等棘手问题，进而达成最佳的算力性能表现与网络利用率。

(2) 由中国移动发挥牵头作用，联合多家中国厂商共同提出的全调度以太网（GSE）。此方案原创提出新型以太网转发和调度机制，将工作重点置于解决网络拥塞以及负载均衡等关键问题上。

在打造新一代网络的架构演进中，GSE技术已进入实践阶段。自2023年5月白皮书首次提出GSE概念以来，通过行业界的通力合作，该技术在标准制定、设备研发等关键领域均取得显著突破。

### 3 GSE 网络的实践

#### 3.1 GSE 关键技术

为应对智算网络流量特性所引发的网络拥塞难题，GSE引入两项关键技术：基于包容器的负载均衡技术以及端到端拥塞避免技术。

##### 3.1.1 基于包容器的负载均衡技术

在负载均衡策略的抉择上，GSE采用喷

洒技术，并提出以等长包容器作为喷洒单位的方案。这一举措不仅确保了负载分担的均匀性，显著降低了数据传输中的乱序程度，减小了后续保序操作所需付出的代价。

由于以太网支持变长包长传输，若仅采用简单的逐包喷洒方式，当包长可变时，极易导致负载分担不均衡。为化解因变长包引发的喷洒分担不均难题，业内运用切包与拼包两项技术加以应对。

切包技术（如图4所示）是指把数据包切割成等长信元后进行喷洒操作，在目的端再对信元进行重组以恢复数据包并实现转发。不过，该技术既要执行信元切分，又要进行组包，实现过程颇为复杂，并且每个信元都需要额外添加信元头，这无疑会造成较大的带宽开销。

拼包技术的运作机制是，将多个数据包组合成等大小的聚合帧后进行喷洒传输，在目的端对聚合帧解帧，还原出原始数据包并予以转发，具体过程如图5所示。相较于切包技

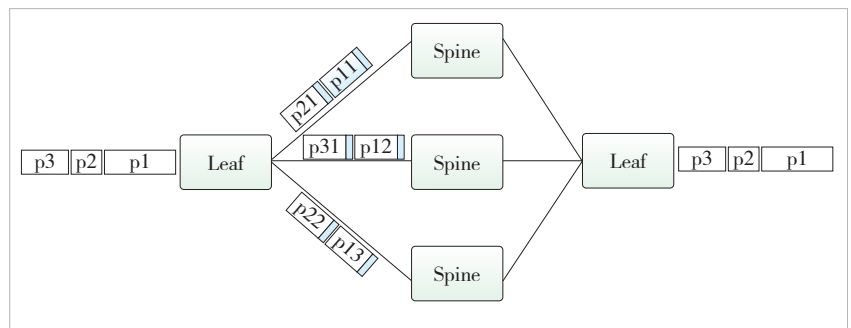


图4 切包喷洒示意图



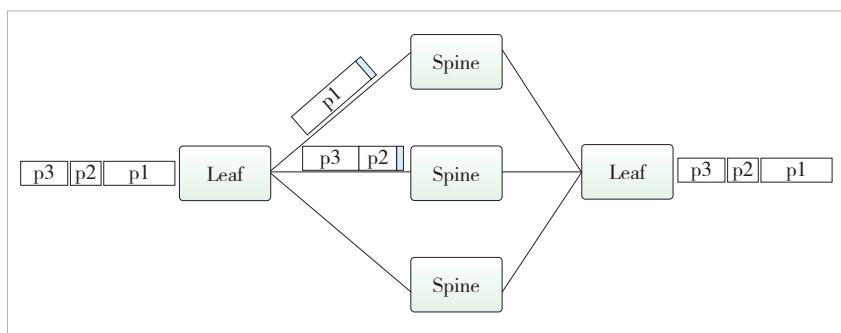


图5 拼包喷洒示意图

术, 拼包技术在实现上更为简易, 多个数据包只需共用一个聚合帧头, 这样可以显著降低带宽开销。然而, 拼包过程会引入不确定的等待时长, 这就导致转发时延抖动情况难以确定。在对同步性要求较高的智算场景中, 这种不确定性会对整体性能产生负面影响。

GSE融合了逐包喷洒低时延以及拼包交换高均衡性的优势, 引入了基于报文容器(PKTC)的转发与负载均衡机制, 具体如图6所示。该机制会把发往相同目的地的数据包, 整合组装成“定长”的虚拟容器来进行转发操作。在这一过程中, 同一容器内的数据包均被标记相同标识, 确保它们沿着相同路径转发, 以此实现保序传输。在进行负载均衡调度时, 该机制以报文容器作为分担单位。值得注意的是, 这里的报文容器属于逻辑概念, 并非是实体, 因而在对数据包进行组装与还原的过程中, 无需额外的硬件投入。每个数据包仅需添加报文容器标识即可。与切包喷洒技术中每个信元都要添加信元头的做法相比, 该方法极大地降低了带宽损耗, 达到更好的效果。

我们假设网络中有 $N$ 条等价路径, 流量的包长在 $[64 \text{ B}, 1500 \text{ B}]$ 之间随机分布, 需要分担的总流量大小为 $F_{\text{total}}$ 字节, 路径的平均负载为 $L_{\text{avg}}$ , 则 $L_{\text{avg}} = \frac{F_{\text{total}}}{N}$ 。在传统逐包喷洒中, 路径 $i$ 上的负载 $L_i$ 与其承载的包长总和成正比, 即 $L_i =$

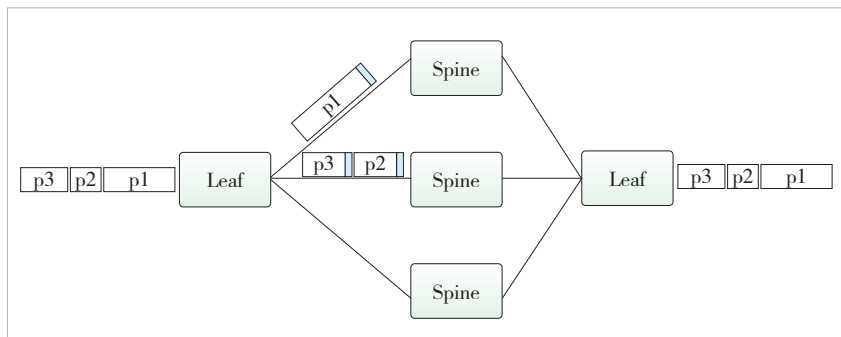


图6 容器喷洒示意图

$\sum_{k=1}^{M_i} P_k$ 。其中 $M_i$ 为路径 $i$ 上的包数,  $P_k$ 为随机变量表示包长。总体方差 $\text{VAR}(L) = \frac{\sum_{i=1}^N (L_i - L_{\text{avg}})^2}{N}$ , 由于 $P_k$ 的变异性,  $L$ 的方差较大, 导致负载分布不均。采用PKTC技术后, 每个容器被设计为等长 $C$ , 路径 $i$ 上的负载为 $L_i = C \times K_i$ , 其中,  $K_i$ 为该路径上的容器数。由于容器可均匀装载, 则 $K_i \approx \frac{F_{\text{total}}}{N \times C}$ ,

那么 $L_i \approx \frac{F_{\text{total}}}{N}$ 。总体方差 $\text{VAR}(L) \approx$

$\frac{\sum_{i=1}^N (\frac{F_{\text{total}}}{N} - \frac{F_{\text{total}}}{N})^2}{N}$ 。此时, 负载方差显著减小, 趋近于零。

PKTC技术确保同一容器内的数据包沿相同路径传输, 有效减少数据乱序情况与时延抖动问题。假设网络每条路径的时延 $D_i$ 服从正态分布 $N(D, \sigma^2)$ , 其中,  $D$ 为平均路径时延,  $\sigma$ 为网络抖动的标准差, 两路径时延差 $D_1 - D_2 \sim N(0, 2\sigma^2)$  (两独立正态分布之差), 标准差为 $\sqrt{2} \cdot \sigma$ 。在传统的逐包喷洒机制里, 每发送完一个报文, 随即更换传输路径。一旦两条路径之间的时延差大于报文的发送间隔, 便会导致乱序现象的产生。尤其是在数据包连续发送的高负载场景下, 报文的发送间隔几乎趋近于0, 乱序的条件可以简化为 $D_1 - D_2 > 0$ :  $P(D_1 - D_2 > 0) = P(Z > 0)$ , 其中 $Z = \frac{D_1 - D_2}{\sqrt{2\sigma^2}} \sim N(1, 0)$ , 对于正态分布 $N(0, 2\sigma^2)$ , 均值为0,  $P(D_1 - D_2 > 0) = 0.5$ 。采用PKTC技术后, 同一容器内的数据包到达顺序始终保持一致, 从根本上杜绝了容器内部的乱序情况, 乱序概率降为0。不过, 在不同容器之间, 前一个容器的最后一个报文与下一个容器的第一个报文之间, 存在出现乱序的可能性, 而这一乱序概率与传统逐包喷洒技术的乱序概率相同。假设容器平均容纳 $m$ 个包, 则两个连续的容器

一共 $2m$ 个包, 乱序的概率为 $\frac{0.5}{2m}$ , 相比传统

逐包喷洒显著下降。借助定长容器的设计, GSE从理论层面达成了负载在多路径间的均衡分配。这一成果有效提升了网络吞吐量, 优化了资源利用率, 为智算网络的高效稳定运行筑牢根基。

### 3.1.2 端到端拥塞避免技术

为有效应对Incast流量拥塞问题, GSE采

取了发送方预请求机制。发送数据前，发送方需先向接收端请求发送权限，接收端则依据自身接收能力，向发送方授予相应信用（即授权）。这一机制确保发送方的数据发送量不会超出接收端的接收能力，以此实现网络拥塞的有效避免。

假设接收端的处理能力为  $R$ ，发送方为  $S_1, S_2, \dots, S_n$  的发送速率为  $r_1, r_2, \dots, r_n$ 。在传统网络中，若  $\sum_{i=1}^n r_i > R$ ，就会发生拥塞。采用 GSE 技术后，接收端根据自身能力授予总信用  $C \leq R$ ，并分配给各发送方  $c_i$ ，满足  $\sum_{i=1}^n c_i \leq C \leq R$ 。因此，发送速率被限制为  $\sum_{i=1}^n r_i \leq R$ ，这在理论上避免了拥塞的发生。

GSE 基于图形处理器（GPU）间实际流量状况，动态构建虚队列，以此减少网络设备所需的队列数量资源。虚队列的调度依据接收端所授予的权限来执行。如此一来，来自不同源设备、发往同一目的端口的多个虚队列，能够依据目的端口的发送能力，进行统一的发送调度，进而达成整个网络的全局调度效果。GSE 将这一技术命名为动态全调度队列（DGSQ）。借助 DGSQ 技术，GSE 网络得以实现端到端的拥塞避免，具体原理如图 7 所示。

当网络出现 Incast 拥塞时，该方案通过将拥塞流量分布式缓存在参与传输的多个源端设备上，实现了全网缓存资源的协同利用。相较于传统方案仅能在目的端设备缓解 Incast 突发流量，这种分布式缓存机制使得网络整体缓存效率得到数量级提升，从而更有效地吸收突发流量。采用 DGSQ 技术后，当源设备本地缓存达到预设阈值时，系统会通过本地 PFC 机制直接通知源 GPU 降速。这一方案避免了传统方法中 PFC 信号需从目的端经 Leaf-Spine 网络反向传输的问题，从而有效防止了 Fabric 网络中可能引发的 PFC 风暴及其导致的网络性能骤降。

### 3.2 GSE 性能测试

GSE 基于包容器的负载均衡和端到端拥塞避免两种关键技术，实现了多路径间的均匀负载分担，减少了乱序和时延抖动，提升了网络吞吐量和时延稳定

性。同时，从源头避免了拥塞，提升了缓存效率和拥塞响应速度。理论分析表明，该架构能显著提升网络吞吐量与时延稳定性，尤其适用于智算网络中高带宽、高突发及高并发的流量场景。

我们在实验室搭建了 32 张 GPU 卡的测试环境（每卡配备 200G 端口），在相同模型和组网拓扑下对比测试了 GSE 与 RoCE 网络的性能。如图 8 所示，测试网络分别由 GSE 原型设备和 RoCE 交换机构建。针对 LLama2-13B 大模型的测试表明（如表 1 所示）：在单任务、多任务及 Leaf 上行链路故障 3 种场景中，GSE 网络性能均显著优于 RoCE，平均提升达 47.7%。

### 3.3 GSE 设备实现

以太网交换机的核心功能是数据报文转发。经过多年发展，主流厂商的转发架构已趋于标准化，通常包括接口模

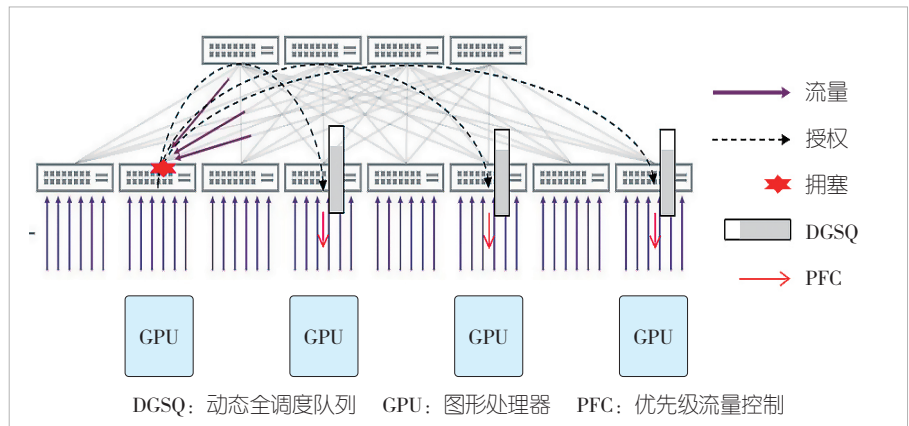


图7 端到端拥塞避免

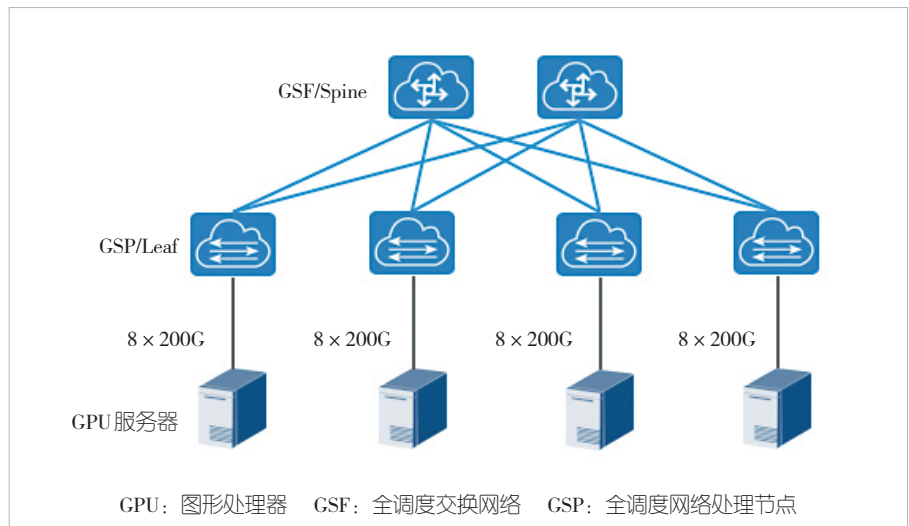


图8 全调度以太网和基于以太网融合远程直接内存访问技术网络的性能测试拓扑



表1 LLama2-13B大模型测试中RoCE网络与GSE网络测试性能数据对比

	RoCE网络性能/ (samples·s <sup>-1</sup> )	GSE网络性能/ (samples·s <sup>-1</sup> )	提升比例/%
单任务	2.66	3.91	47
链路故障(断1条)	2.43	3.84	58
链路故障(断4条)	2.43	3.53	46
多任务	2.45	3.43	40

GSE: 全调度以太网 RoCE: 远程直接内存访问

块、转发引擎、流量管理(TM)单元及辅助功能模块(如内嵌中央处理器、操作维护管理引擎等)<sup>[7-8]</sup>。

GSE在传统以太网交换机的架构之上,融入基于容器的负载均衡技术以及端到端拥塞避免技术。在接口、转发引擎以及TM子系统中增添特定功能特性,可以实现GSE设备的构建,具体实现架构如图9所示。

在设备的接口子系统中,我们增设链路质量监控与故障通告机制。该机制对接口上的前向纠错(FEC)误码情况展开持续监测,以此实现对链路质量的精准预测,确保在链路实际发生故障前,便能敏锐感知潜在问题。一旦监测到异常,立即以物理层编码的形式主动发出故障通告,让全网的GSE设备能够迅速察觉本地及远端故障,实时掌握端到端的可用带宽信息。随后,这些关键信息会及时反馈至授权和容器选路模块,将因链路故障导致的业务中断时长有效控制在了微秒级,大幅提升整个网络的健壮性。

为增强设备的功能与适用性,在其转发引擎里新增对GSE头的解析、封装及转发能力。此新增功能的转发操作将

严格遵循GSE规范要求,全面支持GSE开放生态。

在TM子系统中,我们将队列分配机制从静态分配升级为动态分配,并引入授权管理机制。通过实现队列和缓存资源的池化管理,系统能够根据实际流量需求动态分配资源,仅为活跃流分配队列和缓存。这种优化方案显著减少了大规模组网所需的队列数量,同时大幅提升了整网队列和缓存资源的利用效率。

在TM子系统中增加容器构建与以容器为单位的负载均衡选路功能。系统通过系统通过累计算出队报文的包长,除以配置的容器长度,得到商值作为报文所属容器的ID,并以该容器ID为索引选择转发路径。

在TM子系统中增加容器保序功能,针对从远端接收到的、目的为本地直连GPU的流量,按目的地址进行容器级保序处理。该功能可消除因容器间负载均衡引入的报文乱序,避免将乱序传递给目的端而触发不必要的重传,从而提升系统整体性能<sup>[9]</sup>。

## 4 结束语

随着AI时代的到来,行业对适配大规模算力的网络需求日益迫切。当前人们对Scale-Out网络的探索不断深入,各类实践呈现百家争鸣之势,技术收敛与共识形成尚需时日。

其中,面向下一代网络重构的重要技术体系GSE,针对智算网络特有的低熵、大象流、同步突发等流量特征,我们创新性地提出几大核心技术:基于GSE封装头的转发机制、

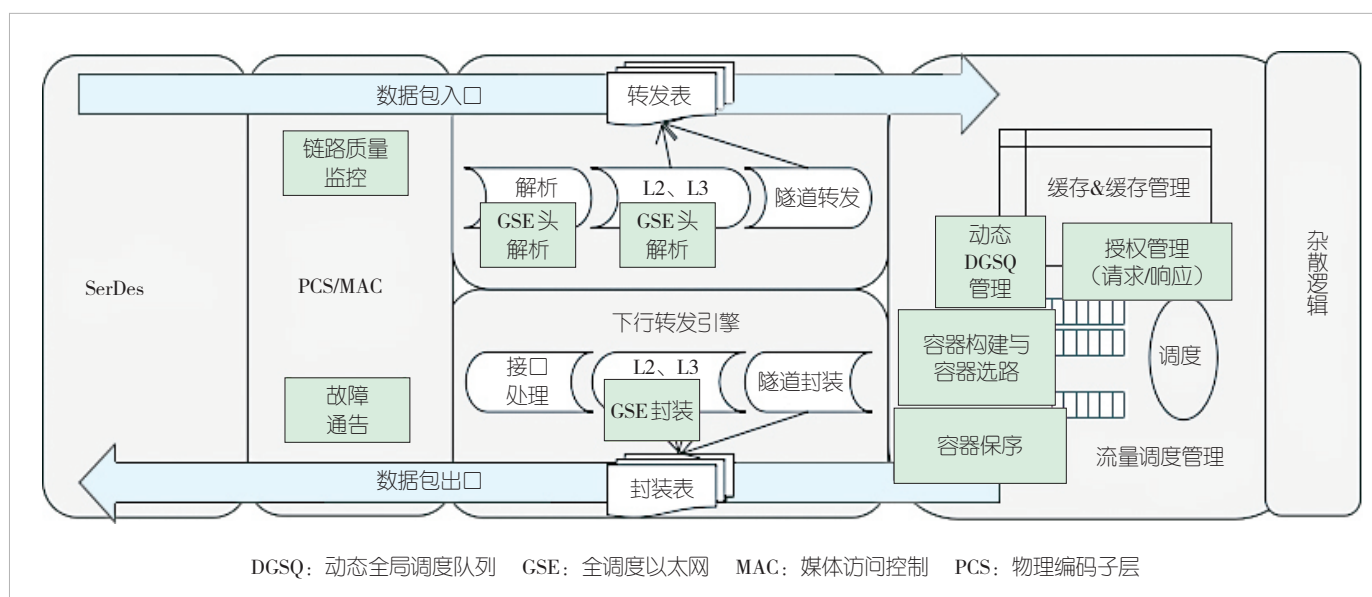


图9 GSE设备新增特性

基于 PKTC 的负载均衡、基于 DGSQ 的授权与调度等。这些技术有效解决了当前智算 Scale-Out 网络面临的诸多痛点：传统基于流的等价多路径负载均衡效果不佳，大规模 Incast 流量易引发拥塞甚至丢包等问题。当前，GSE 相关硬件研发、协议创新及工程实践已进入快速发展阶段，多项核心技术通过原型机验证取得突破性进展，展现出卓越的性能表现和产业化潜力，正成为智能计算 Scale-Out 网络架构最具发展前景的技术路线之一。

#### 参考文献

- [1] 中国移动通信研究院. GSE 2.0 网络侧优化技术标准 [R]. 2024
- [2] JIANG Z H, LIN H B, ZHONG Y M, et al. MegaScale: scaling large language model training to more than 10, 000 GPUs [EB/OL]. [2025-03-02]. <https://arxiv.org/abs/2402.15627>
- [3] QIAN K, XI Y Q, CAO J M, et al. Alibaba HPN: a data center network for large language model training[C]//Proc of ACM SIGCOMM. ACM, 2024: 691-706
- [4] 中国移动通信研究院. 全调度以太网技术架构白皮书 [R]. 2023
- [5] 中国移动通信研究院. 面向 AI 大模型的智算中心网络演进白皮书 [R]. 2023
- [6] WANG S, GAO K H, QIAN K, et al. Predictable vFabric on informative data plane [EB/OL]. [2025-03-02]. <https://dblp.org/rec/conf/sigcomm/0028GQ0MLZZSGZF22.html>
- [7] 段晓东, 程伟强, 王瑞雪, 等. 面向新型智能计算中心的全调度以太网技术 [J]. 中兴通讯技术, 2023, 29(4): 57-63. DOI: 10.12142/ZTETJ.202304011
- [8] 段晓东, 陆璐, 孙滔, 等. 广域抗损高吞吐 URDMA 技术 [J]. 中兴通讯技术, 中兴通讯技术, 2024, 30(6): 23-30. DOI: 10.12142/ZTETJ.202406005
- [9] 崔佳怡, 谢人超, 唐琴琴. 基于生成式人工智能的算力网络自智优化研究综述 [J]. 中兴通讯技术, 2024, 30(6): 54-62. DOI: 10.12142/ZTETJ.202406009

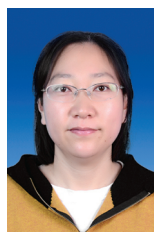
#### 作者简介



**程伟强**，东南大学网络空间安全学院在读博士，中国移动通信有限公司研究院基础网络技术研究所副所长，教授级高工；主要从事下一代互联网、数据中心网络、传输网等方面的技术研究和标准推动工作；参与 IETF、ITU-T 等 10 余项国际标准的制定。



**李新双**，中兴通讯股份有限公司承载网产品副总经理；在数据通信产品等相关领域拥有 20 余年工作经验。



**白艳**，中国移动通信有限公司研究院基础网络研究所项目经理；主要从事数据中心网络技术与方案研究工作。



**吕勇**，中兴通讯股份有限公司有线系统架构师；主要从事数据通信网络设备架构等的研究工作。

# 基于大语言模型蒙特卡洛树搜索的智算网络故障根因分析系统



## Fault Root Cause Analysis System of Intelligent Computing Networks Based on Large Language Models and Monte Carlo Tree Search

罗子秋/LUO Ziqiu, 苗宇锴/MIAO Yukai, 李丹/LI Dan

(清华大学, 中国 北京 100080)  
(Tsinghua University, Beijing 100080, China)

DOI: 10.12142/ZTETJ.202502004

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250424.0902.002.html>

网络出版日期: 2025-04-24

收稿日期: 2025-02-20

**摘要:** 提出了一种基于大语言模型 (LLM) 进行蒙特卡洛树搜索的智算网络故障根因分析系统 (RCA-MCTS)。利用 LLM 推理研究领域在蒙特卡洛树搜索上的前沿研究, 面向智算网络复杂故障场景, 设计了适用于故障根因分析任务的多策略提示语扩展机制, 并基于与故障模拟环境交互反馈的方式设计了模拟机制, 使得 LLM 推理时的蒙特卡洛树搜索过程适配于故障根因分析任务场景。实验表明, RCA-MCTS 在故障根因分析任务准确率上提升 33%~43%, 在故障推理动作序列平均匹配度上提升 18%~34%。

**关键词:** 智算网络; 故障根因分析; 大语言模型; 蒙特卡洛树搜索

**Abstract:** A fault root cause analysis (RCA) system of intelligent computing networks based on Monte Carlo tree search (MCTS) and large language models (LLM), named RCA-MCTS, is proposed in this paper. By leveraging cutting-edge research on MCTS in the domain of LLM reasoning, a multi-strategy prompt expansion mechanism is designed for fault root cause analysis tasks in intelligent network fault scenarios. Additionally, a simulation mechanism is developed based on feedback interactions with the fault environment, enabling the MCTS process during LLM reasoning to be adapted to the fault root cause analysis task. Experimental results show that RCA-MCTS improves the accuracy of fault root cause analysis by 33%–43%, and enhances the average matching degree of fault inference action sequences by 18%–34%.

**Keywords:** intelligent computing network; fault root cause analysis; large language model; Monte Carlo tree search

**引用格式:** 罗子秋, 苗宇锴, 李丹. 基于大语言模型蒙特卡洛树搜索的智算网络故障根因分析系统 [J]. 中兴通讯技术, 2025, 31(2):21–30. DOI: 10.12142/ZTETJ.202502004

**Citation:** LUO Z Q, MIAO Y K, LI D. Fault root cause analysis system of intelligent computing networks based on large language models and Monte Carlo tree search [J]. ZTE technology journal, 2025, 31(2): 21–30. DOI: 10.12142/ZTETJ.202502004

为了支撑日益增长的通信和计算需求, 现代网络系统不断向更庞大的规模、更复杂的架构演进。自 2022 年 ChatGPT<sup>[1]</sup> 发布后, 大语言模型 (LLM) 的应用成为热门话题, 由此带来的算力需求爆炸式地增长, 也对支撑 LLM 训练和推理的网络系统提出了更高的要求。人工智能算力网络 (简称智算网络) 是一种专门服务于大规模 LLM 训练和推理的新型网络系统, 提供了灵活调度、资源共享、统一服务的能力。智算网络往往具有复杂的拓扑设计, 集成了大量异构的新型算力设施, 这给网络运维带来了新的挑战。目前智算网络的运维主要依赖于传统网络运维的专家经验, 可能存在对新型算力设施的运维经验不足、不同技术背景的运维人员

之间协调效率低等问题。面对可能存在的故障, 发现不及时、响应慢、解决不彻底, 将会对 LLM 的训练和推理任务产生严重的影响。

针对传统人工运维方法的不足, 学术界已提出一系列基于人工智能的运维 (AIOps) 方法, 可应用于故障检测、故障定位、故障根因分析 (RCA)、故障恢复等多种任务<sup>[2–3]</sup>。本工作聚焦于故障根因分析任务, 研究基于人工智能模型分析故障信息和环境数据, 输出可能的故障根因, 指导运维决策。近年来, 由于 LLM 在理解和生成自然语言以及执行复杂的推理任务上表现出了超越以往智能模型的卓越能力, 学术界已有将 LLM 应用于 AIOps 的工作。例如, RCA Copilot 使

用 LLM 生成结构化告警数据的文本摘要，并根据语义相似度从历史告警的故障根因中检索出可能的根因<sup>[4]</sup>；RCAgent 利用 LLM 生成与故障告警信息高度相关的根因分析等内容，在生成过程中通过多路径并行推理的一致性检查，提升输出内容的可靠性<sup>[5]</sup>。然而，这些方法完全依赖 LLM 的领域知识，将各项结构化告警信息以提示语的形式输入 LLM 后，只通过单步推理就直接获取模型预测的故障根因，这与网络系统故障运维的实际情况存在较大差异。在实际故障运维场景中，运维人员需要根据故障处理经验与环境进行多轮交互来确定故障根因。在每一轮交互中，运维人员要根据故障信息和当前状态，决定下一步要执行的运维动作，随后执行此动作，并获取环境反馈，以帮助下一轮的决策。因此，我们认为故障根因应由故障告警信息和若干<运维动作，环境反馈>二元组来确定，即： $RC = F(Alert, (Action, Observation) \times N)$ ，其中  $N$  需要足够大，以唯一地确定故障根因。

根据上述讨论，如图 1 所示，我们提出一个基于 LLM 的自动与环境进行多轮交互并分析故障根因的智能体。该智能体以 LLM 为核心，配备智算网络运维知识库，在获得输入的故障告警信息后，通过多次调用运维工具与故障环境进行交互，最终定位故障根因。

RCA 任务对 LLM 的能力提出了很高的要求：LLM 不仅要具备基本的知识问答能力，还要具备复杂领域知识的理解

能力和多步迭代的推理能力。现有的很多 LLM 尽管在各类通用任务上具备强大的能力，但在完成复杂推理任务时仍表现不佳。例如，Mistral-7B 等主流 LLM 在使用链式思维等技术<sup>[6-8]</sup>增强推理能力的情况下，在 GSM8K 等需要多步推理的数据集上的最高准确率仅为 36.5%。计划推理（RAP）<sup>[9]</sup>等改进的链式推理框架采用了一种自我探索的方案，通过自我奖励反馈，迭代地提高 LLM 的推理表现。但是，这种推理方式难以有效地探索解空间，即使经过多次尝试，也常常困在一个低质量的区域中。LLM 多步推理能力的不足显著阻碍了 RCA 智能体的开发和应用。

近段时间以来，以 OpenAI o1 模型为代表的 LLM 多策略推理技术获得了较大的进展，为 RCA 智能体推理能力的实现提供了思路。面对 GPT-4、GPT-4o 等模型都难以回答的奥林匹克数学竞赛题目，o1 借鉴人类思考方式，自主优化思考过程，尝试多种策略，识别思考过程中的错误，最终完成复杂逻辑计算的推理。相比于链式推理以及基于链式推理数据的微调，这种推理方式能够探索更大范围的解空间，提高推理出最终正确结果的概率。不过 OpenAI 并没有公开 o1 推理框架的具体实现方法，一系列复现 o1 的开源研究<sup>[10-17]</sup>大多使用蒙特卡洛树搜索（MCTS）这一经典算法来实现类似 o1 的推理框架。基于 MCTS 算法，即使在模型基础能力不足、数据集质量不佳的情况下，也能够增强 LLM 的推理能

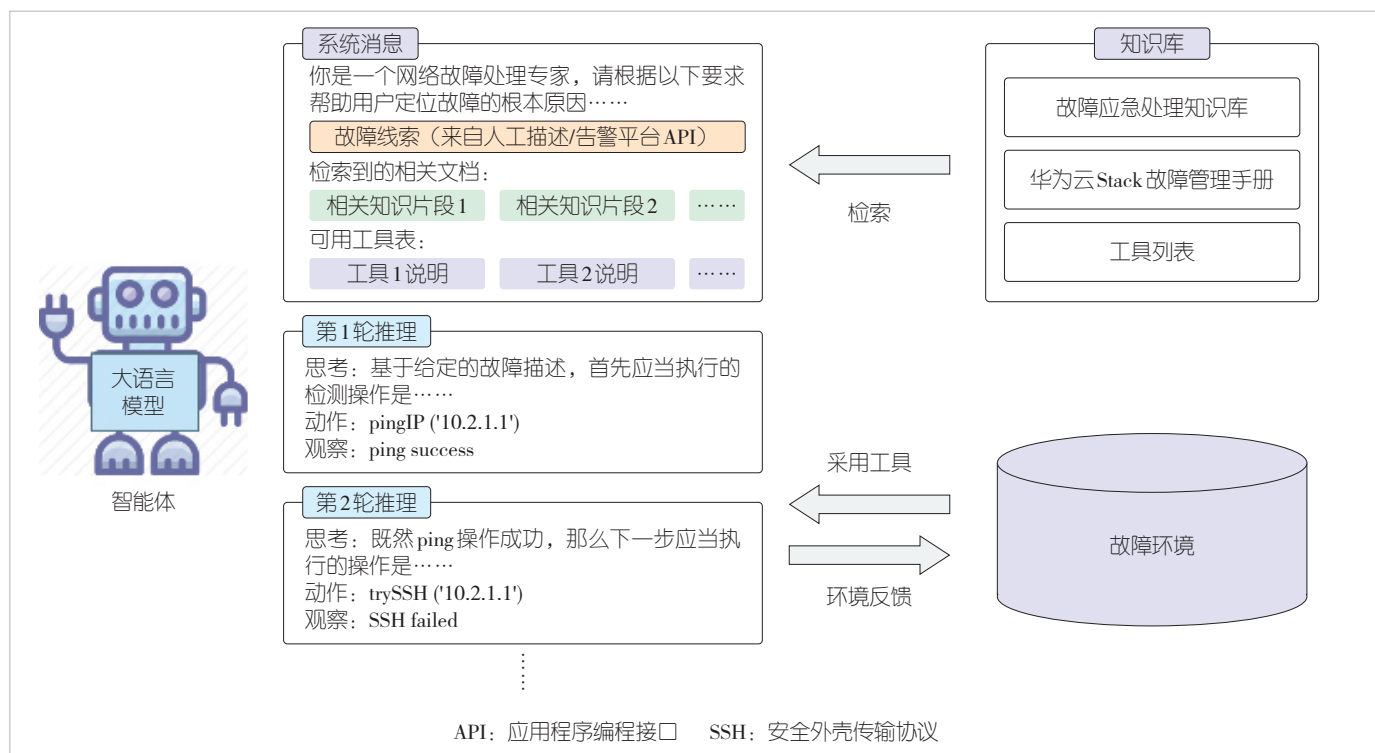


图1 故障根因推理智能体



力。但是这些工作的应用场景多为求解数学问题，与需要多轮环境交互的RCA场景有很大区别，直接套用现有的MCTS推理框架并不可行。针对上述挑战，本文提出了适用于RCA任务的MCTS推理框架RCA-MCTS，通过设计符合RCA场景特征的MCTS扩展策略和环境交互机制，显著提升了7B参数规模的LLM在RCA任务上的端到端表现。

## 1 RCA-MCTS研究动机

MCTS核心思想是将统计模拟（Monte Carlo方法）和树搜索（Tree Search）相结合，平衡“探索（Exploration）”与“利用（Exploitation）”，逐步建立和扩展决策树，并覆盖多路高价值的决策路径。MCTS以模型初始输入信息作为根节点，迭代式地进行多轮计算，每轮过程包含选择、扩展、模拟、回传4个步骤。在选择阶段，算法从根节点出发，根据平衡探索与利用的策略选择出一条路径，路径的终点作为当前节点；在扩展阶段，算法在当前节点基于动作扩展策略添加一个新的子节点；在模拟阶段，算法以扩展的新节点为起点进行多次模拟过程，每次模拟过程通过一步步的随机采样探索出一条新的子路径，直到达到预设的终止状态，并根据终点的正确性获得奖励分数；在回传阶段，算法将多次模拟的奖励回传给路径上的所有节点，更新节点的价值，并以此影响下一轮的选择过程。在现有的基于MCTS的LLM多策略推理研究中，MCTS搜索树的每一条连接父子节点的边对应LLM的一步推理，多种提示语策略构成了MCTS扩展阶段的动作空间，LLM在MCTS的指引下提高了生成正确答案的概率。

然而，在上述基于MCTS的LLM推理研究中，绝大多数选择求解数学问题作为任务场景，这与RCA任务存在如下差异：第一，数学题目的解决过程只需要根据推理历史选择合适的推理策略，现有方案的MCTS扩展仅需遵循静态的、预定义的规则。而RCA任务则需要考虑环境的反馈信息，即使在推理历史相同的情况下，故障环境反馈仍存在多种可能，对于不同的环境反馈信息，应采取不同的故障分析策略。因此，LLM需要对推理历史和环境交互历史信息进行综合处理，不断动态调整策略。第二，诸如GSM8K的数学题目数据集中，每道题均存在唯一的、正确的数值解，在MCTS模拟评估阶段可以直接基于规则提取数值答案，通过数值计算、置信度评估、一致性检查或基于外部LLM判断等方式，验证答案的正确性，并以此计算奖励分数。而RCA任务的最终输出为文本形式的故障根因描述，不能直接利用数值计算或置信度评估来验证正确性，一致性检查或者基于外部LLM判断则需要评判模型是否具有RCA领域专

业知识和推理能力。第三，主流的LLM的预训练数据集普遍包含了海量的数学相关数据，这使得LLM本身已具备较强的数学基础能力，可通过MCTS推理框架更进一步增强数学推理能力。而RCA任务的专业性相对强得多，且LLM预训练过程中缺少海量高质量的RCA数据集，导致模型本身RCA基础能力较弱。MCTS推理框架虽然能在一定程度上提升推理表现，但绝对的端到端准确率仍然较低。

基于上述讨论，我们列出实现RCA-MCTS的主要挑战：

- 1) 故障环境反馈的不确定性。我们需要针对可变的环境反馈，设计MCTS扩展阶段的推理策略。
- 2) 故障推理中奖励分数计算的复杂性。我们需要设计有效的奖励函数，为模拟阶段产生的推理节点计算奖励分数。
- 3) LLM故障运维基础能力的不足。我们需要提升LLM在RCA任务领域的基础能力，从而进一步提升使用MCTS方法后的端到端效果。

针对上述挑战，本文设计了RCA-MCTS故障根因推理框架，使得LLM在RCA任务场景下进行多策略并行式推理，输出准确的故障根因。RCA-MCTS在MCTS扩展阶段运用适配RCA场景的提示语策略，综合分析推理历史数据和故障环境反馈信息，合理规划下一步推理行动。在模拟阶段，RCA-MCTS结合故障环境交互和我们基于过程标注数据训练的过程奖励模型（PRM）计算推理节点的奖励分数，从而引导MCTS的路径搜索。我们构建了一个故障向量知识库，为LLM推理提供故障运维示例，并基于故障告警信息设计故障环境，为LLM采取的故障运维动作提供唯一确定的故障反馈。模拟过程中，RCA-MCTS支持LLM与故障环境持续交互，直至输出根因或者达到指定的最大搜索深度，基于故障知识库对输出根因进行验证。同时，受到AlphaGo工作的启发<sup>[18]</sup>，我们训练了PRM模型，用于评估MCTS中间节点对终局推理结果的贡献价值，以便提供更可靠的奖励反馈。最后为了进一步提升模型在推理框架下的故障推理能力，我们构建了ReAct格式的标准推理数据集，以此对模型进行了微调训练。微调后的模型增强了RCA领域的基础能力，结合RCA-MCTS推理框架展现出更强的故障根因分析能力。

结合以上所述，本研究的主要贡献包括：

- 1) 构建RCA-MCTS推理框架，针对故障环境反馈设计MCTS动态扩展策略。
- 2) 构建故障知识库、故障模拟环境，作为RCA-MCTS推理框架的外部依赖，支持LLM与环境的交互、输出根因的验证评估，模拟现实运维场景，客观评估RCA-MCTS推理框架性能。
- 3) 训练面向RCA任务奖励计算的PRM模型，用于辅助

MCTS模拟评估。

4) 基于自动化流程构造 ReAct 格式故障根因推理数据集, 训练核心 LLM, 提升模型故障根因推理基础能力。

5) 通过实验证明了 RCA-MCTS 故障根因推理框架的效果。以 Qwen2-7B-Instruct 模型为基座模型时, RCA-MCTS 在故障根因推理数据评测集的端到端根因分析准确率、动作决策序列合理匹配度两个指标上分别取得了 43.6% 和 34.0% 的提升。对于其他 7B 模型, 我们的方法也显著提升了推理表现。

本文将在第 2 节介绍 RCA-MCTS 故障根因推理框架的整体架构, 在接下来的第 3 节重点论述推理框架的扩展策略和模拟过程设计以及相关模型的训练, 并在第 4 节中给出实验设置、实验结果和结果分析。

## 2 RCA-MCTS 推理框架架构设计

如图 2 所示, RCA-MCTS 推理框架以原始故障告警信息作为输入, 基于语义相似度在故障运维知识库中匹配可采取的故障运维动作集合、待确定的故障根因范围; 初始提示语模板对告警输入、运维动作集合、可能的根因范围进行文本组合, 并在系统提示语中定义模型推理范式遵守 ReAct 规则以确保模型每步推理的规范性和有效性<sup>[19-20]</sup>, 组合后的文本加上系统提示语构成初始提示语; 推理框架的核心 LLM 接收初始提示语作为输入, 在基于面向 RCA 任务场景设计的 MCTS 扩展策略指导下, 综合分析推理历史和故障环境反馈

信息, 合理选择提示语模版进行下一步推理; RCA-MCTS 基于推理内容在故障知识库提供的运维动作集合中匹配运维动作, 并采取运维动作与故障模拟环境交互获得反馈信息并记录, 产生新的 MCTS 推理节点; 随后, RCA-MCTS 基于新节点按照规则与故障模拟环境模拟交互至终局状态, 并结合 PRM 完成节点评估, 回传节点分数, 完成 MCTS 一轮过程; 在进行多轮 MCTS 构建过程后, 核心 LLM 输出故障告警对应的 RCA 推理树, 对终节点的各项属性进行综合分析, 基于多数投票的方式决定唯一的故障根因作为推理框架的最终输出。

在进行故障根因推理之前, RCA-MCTS 预先构建一个故障运维知识库和一个故障模拟环境作为依赖, 前者用于提供推理框架下故障处理可使用的动作描述集合和待分析的故障根因范围, 后者用于接收模型推理过程中采取的运维动作, 提供环境反馈信号。对于每一条故障告警输入, 故障模拟环境对于故障运维动作集合提供的反馈信号是唯一确定的。若干<故障运维动作, 故障环境反馈>信息对结合原始故障告警输入, 对应唯一的故障根因。核心 LLM 接受原始故障告警后, 基于 RCA-MCTS 推理框架不断通过推理获取动作和环境反馈信息对, 最终推理出故障根因。

## 3 RCA-MCTS 方法

### 3.1 RCA-MCTS 外部依赖构建

RCA-MCTS 外部依赖包含故障运维知识库和故障模拟

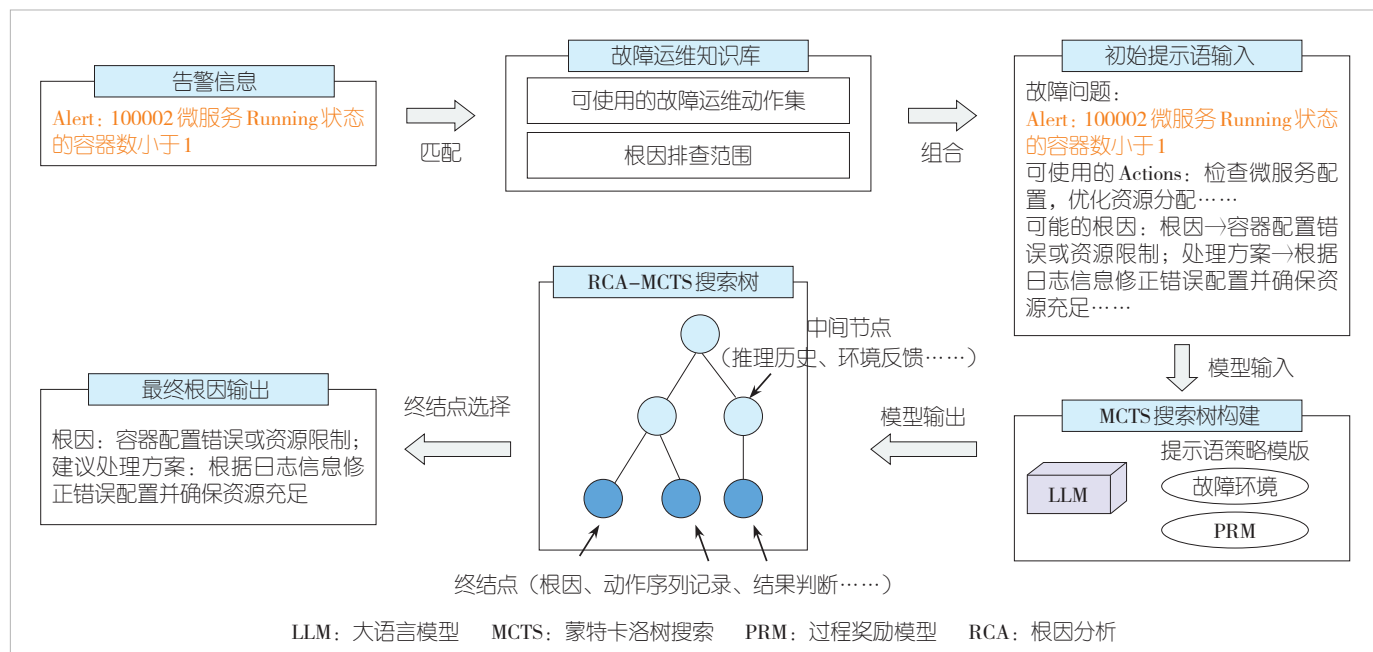


图2 RCA-MCTS推理框架整体架构



环境。故障运维知识库提供可使用的运维动作描述和候选的根因集合范围，并将它们作为初始输入的系统提示语部分；故障模拟环境在MCTS扩展阶段对核心LLM采取的运维动作进行反馈，在MCTS模拟评估阶段对最终输出的根因进行验证评估。二者均是完成完整MCTS推理过程的关键组件，需要在执行MCTS算法之前构建。

为了构建故障运维知识库，我们通过工具爬取了17个PDF格式的华为云故障文档和1个Word格式的云网络环境故障处理知识库，作为构建故障运维知识库的数据源。我们采取一系列文档处理规则对文档内容进行格式清洗和关键内容提取，挖掘出3 203个故障处理文本描述块。接下来，我们使用GPT-4在自设计提示语方法下，将每个故障文本描述块转化成ReAct格式的故障根因推理思维链数据，每条数据包含告警信息、故障根因信息，以及多步故障运维动作和故障模拟环境反馈。一般情况下，同一个故障告警在不同故障模拟环境反馈下对应于不同故障根因。我们进一步对这些数据进行清洗，滤除格式不合规的数据，最终构造了包含1 507个故障告警信息和3 818条ReAct格式的故障根因推理思维链数据集。我们通过对思维链数据的处理步骤、故障根因的整理和向量化，构造了故障运维动作集合和故障根因集合的向量数据库，供RCA-MCTS在开始阶段基于故障告警输入进行检索增强（RAG）以生成系统提示语。

故障运维思维链数据集除了用于构建故障运维知识库，还用于PRM和核心LLM的训练评估以及推理框架端到端评测实验。我们基于该数据集进行采样、划分、预处理工作，这些工作服务于对应的模型训练和评测环节，数据处理的细

节将在后续小节中论述。

在构建故障模拟环境时，基于 $RC = F(Alert, (Action, Observation) \times N)$ 的模式定义，对于每一条故障告警输入，我们在故障模拟环境中为每个运维动作设置了唯一确定的动作反馈，并针对原始故障告警输入设置唯一确定的故障根因作为答案。当核心LLM经过多轮推理达到终局状态时，我们基于文本规则提取出根因答案作为最终根因输出，在可能故障根因范围中匹配根因，随后基于故障知识库提供的验证方法，在故障模拟环境中验证核心LLM推理出的根因的正确性。

### 3.2 面向RCA任务的MCTS过程设计

如图3所示，RCA-MCTS推理过程遵循MCTS算法步骤多次循环，构建故障根因推理搜索树。在选择阶段RCA-MCTS沿用了MCTS经典算法的默认设置，基于“探索-利用”准则递归地选择待扩展的叶子节点，但在MCTS扩展阶段和模拟评估阶段采用了适用于RCA任务的设计。这些设计有助于RCA-MCTS推理框架在故障推理任务场景下不断动态调整策略和并行式推理，在故障推理数据集上的端到端准确率提升方面起到了关键作用。

#### 1) RCA-MCTS扩展过程

RCA-MCTS扩展阶段的策略综合考虑了推理历史和故障模拟环境反馈。RCA-MCTS扩展动作定义为一系列适用于核心LLM进行故障根因推理的提示语，包括以下几类：（1）提示核心LLM对过去推理历史进行归纳总结，提炼当前故障根因推理思路，专注于异常的故障模拟环境反馈信息；（2）提示核心LLM判断故障模拟环境反馈信息是否异

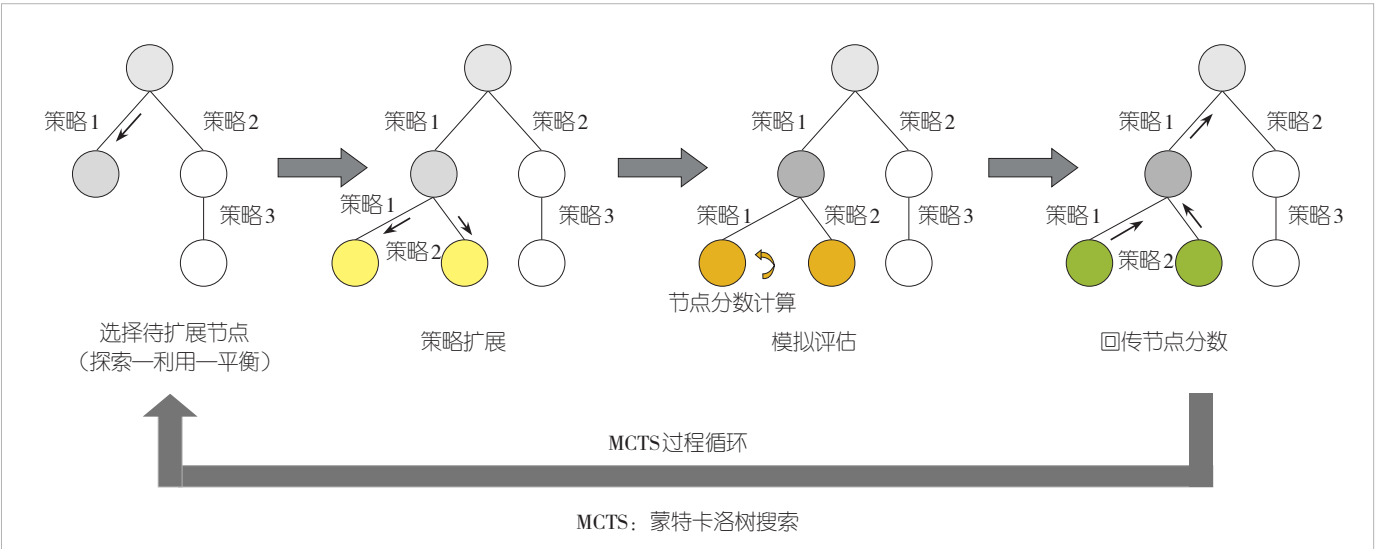


图3 RCA-MCTS搜索树构建过程

常，结合推理历史思考下一步应当采取的动作，以获取更多的环境反馈作为推理依据；(3) 提示核心 LLM 根据过去推理历史，判断故障告警最可能的根本原因；(4) 提示核心模型回顾 ReAct 格式的推理历史，反思采取故障运维动作是否合理、故障模拟环境反馈是否与根因匹配。如图 4 所示，MCTS 节点基于当前节点状态（如推理链条长度、推理历史、过去推理所采取的推理策略等）选择多个提示语策略模版，每个被选择的提示语模版与节点内记录的系统提示语、推理历史组成扩展提示语输入；核心 LLM 接受每个扩展提示语进行单步扩展推理，输出扩展推理内容，扩展产生新的 MCTS 推理节点。特别地，当扩展策略中包含故障运维动作时，RCA-MCTS 提取 LLM 输出的故障运维动作内容片段，随后基于语义相似度，在故障知识库提供的故障运维动作集合中匹配最近似的动作，采用动作与故障模拟环境进行交互并获取环境反馈，构成新的推理历史。基于这种扩展方式，RCA-MCTS 产生多个新的 MCTS 子节点，在多种思维策略方向上并行推理，扩大故障根因搜索空间。随后 RCA-MCTS 基于故障模拟环境多轮交互过程和过程奖励模型，对扩展节点进行模拟评估，为下一轮 MCTS 过程中选择更有价值的节点提供判断依据，提高推理出正确根因的可能性。

## 2) RCA-MCTS 模拟评估过程

RCA-MCTS 模拟评估环节对扩展产生的 MCTS 节点进行评分，如图 5 所示。节点分值由基于故障模拟环境交互过程

的模拟结果评分和 PRM 评分共同决定。其中，故障模拟环境交互模拟是指：从当前节点出发，进一步选择策略进行核心 LLM 推理，与故障模拟环境交互后产生新的 MCTS 扩展节点，在扩展节点中随机选择节点进一步推理产生新节点，如此推理多次直到能从核心 LLM 输出中提取故障根因，或者达到规定的树最大深度。当 MCTS 的扩展策略中包含故障根因推断策略时，RCA-MCTS 从核心 LLM 输出中提取故障根因答案，在故障模拟环境中验证答案的正确性并记录，完成一次故障模拟环境交互模拟。这样的模拟会进行多次，推断出正确根因的模拟次数所占比例即为故障模拟环境交互模拟结果评分。

基于故障模拟环境交互的模拟过程中，由于搜索空间无法穷尽，因此 MCTS 将随机选择节点进行扩展，选用多次模拟的统计结果作为环境交互模拟评分。这种统计采样方法产生的评分具有一定的随机性，特别是在推理成本不足导致模拟次数有限的情况下，分数的不可靠性会被进一步放大。为此，我们引入参数化 PRM 对扩展节点进行价值判断，来平衡模拟统计方法的随机性。在不用耗费大量推理成本增大模拟次数的情况下，PRM 基于当前节点推理历史输入，判断当前核心模型能最终推理出正确根因的概率。将模拟交互得分和 PRM 得分的结果加权相加，最终得到了 RCA-MCTS 模拟评估阶段的节点分数。

## 3) PRM 训练

RCA-MCTS 依赖的 PRM 根据当前推理历史判断已采取

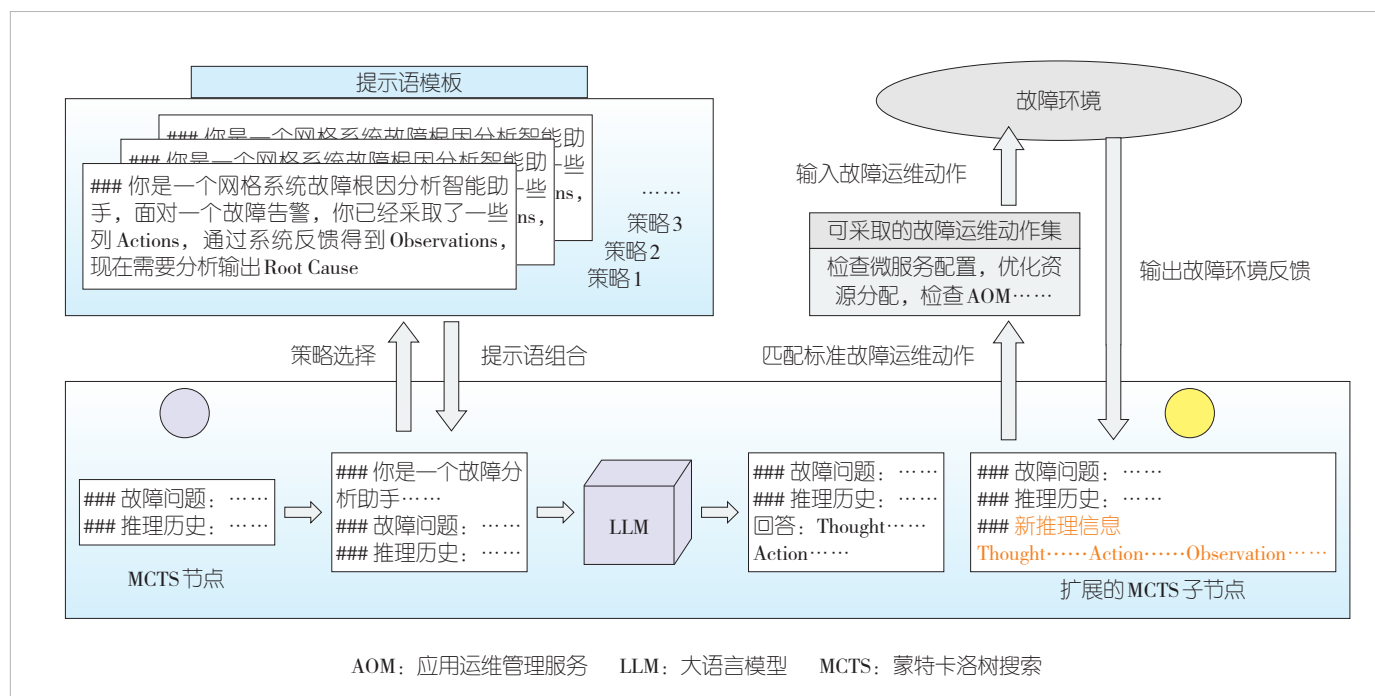


图 4 RCA-MCTS 扩展过程

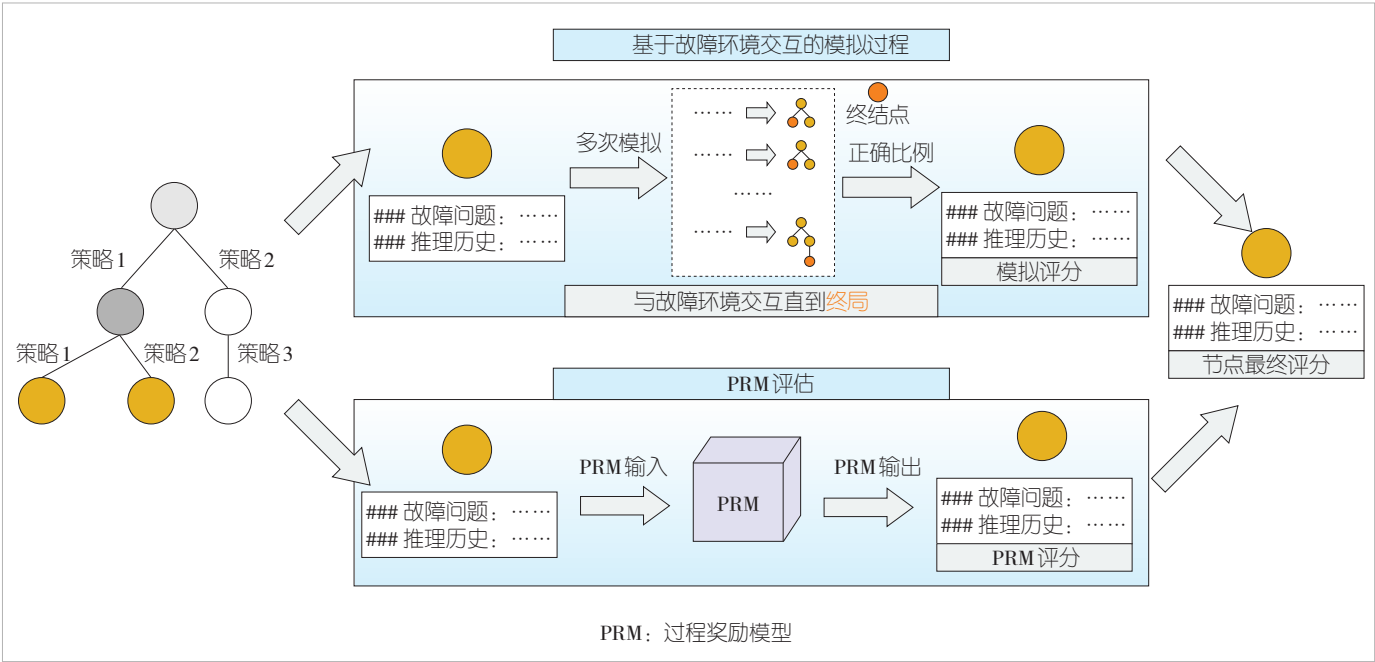


图5 RCA-MCTS模拟评估过程

的故障运维动作的合理性，在不需要 LLM 与故障模拟环境交互至终止状态的情况下，评估当前推理动作最终推理出正确的故障根因的可能性，其输出值的范围为[0,1]连续区间。训练 PRM 依赖于带有推理步骤标签的故障根因推理数据集。为了构建该数据集，我们基于构建故障运维知识库过程中产生的 ReAct 多步推理格式的故障根因推理思维链数据集，从中采样了 500 个故障告警对应的 1 382 条推理数据，通过打乱推理顺序、重复步骤、引入其他告警运维动作等方式，替换掉推理思维链数据中的某些步骤环节，并在错误的推理步骤处标注负标签，在剩余步骤处标注正标签。基于该构造方式，我们获得了包含 6 742 条故障根因推理思维链数据（对应 32 633 个推理步骤标签）的 ReAct 故障根因推理数据集，用于训练 PRM 对每个 RCA-MCTS 中间推理节点的有效评估能力，服务于 RCA-MCTS 的模拟评估过程。

### 3.3 核心模型 RCA 基础能力微调训练

前述开发 LLM 推理框架的工作专注于通过设计多策略推理方法激发模型智能、提高推理效率，并不涉及模型参数微调。在模型故障运维领域基础能力本身缺失的情况下，推理框架能够激发的表现有限，导致 MCTS 搜索树构建过程中获取的正反馈稀疏，降低搜索方法的效率。为了增加与故障模拟环境得到的正反馈，提升 RCA-MCTS 通过高价值推理节点迭代策略的效率，我们进一步使用故障运维专业领域数据集

微调核心 LLM。在构造专业数据集工作上，我们基于构建故障运维知识库中产生的 ReAct 数据集，从中提取了 1 000 条 ReAct 格式的多步故障思维链数据。在训练时，如图 6 所示，

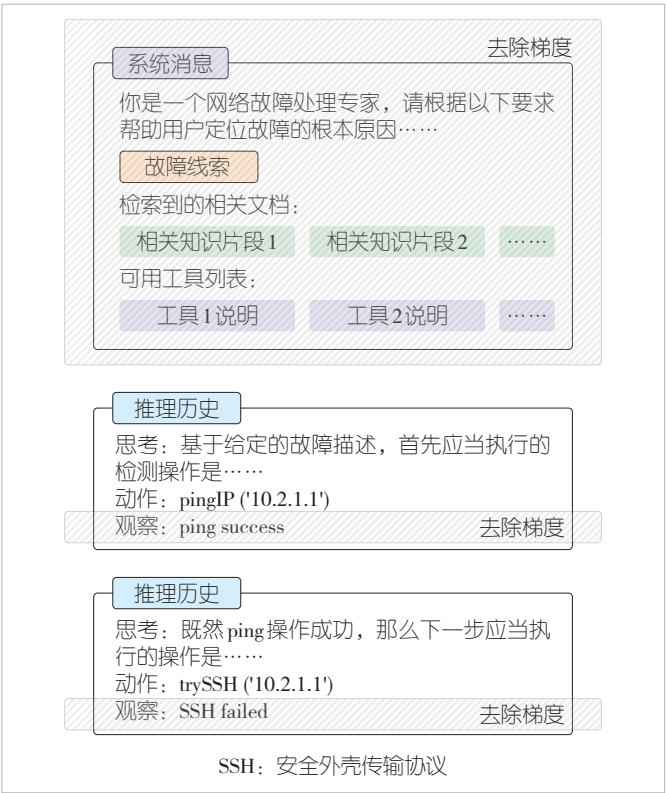


图6 核心模型在故障根因数据集上的训练



我们去除了系统消息前缀和故障反馈环境的梯度，使得核心 LLM 专注于正确的故障根因分析思考和故障运维动作决策，提升了核心 LLM 在故障根因推理专业领域的基础能力。

## 4 RCA-MCTS 评估

### 4.1 评测数据集

我们基于前述从故障运维文档提取到的 ReAct 推理数据集构造评测数据集，用于 RCA-MCTS 推理框架性能端到端评估。该评测数据集包含 300 条故障告警，每条故障告警对应一系列参考运维动作序列和唯一的故障根因。推理框架的目标是输出唯一正确的根因，在推理过程中产生的运维动作序列不必完全遵照参考序列。需要注意的是，该评测数据集与上一节中提到的用于训练 PRM、核心 LLM 的数据集互不包含，以确保评测数据集评估核心 LLM 在推理框架下展现出的故障推理能力，而非单纯对推理数据集进行拟合。

### 4.2 实验设置

我们采用 RCA-MCTS 推理框架在评测集上的故障根因推断准确率，评估其故障根因推断能力。另外，为了评估 RCA-MCTS 在推理过程中采取运维动作的合理性，我们设计了故障推理动作序列匹配度指标。该指标通过计算运维动作中属于评测集参考运维动作序列集的比例获得。我们将 ReAct 方法作为 baseline，对比体现 RCA-MCTS 推理框架的有效性。

在评测模型方面，我们使用 Qwen2-7B-Instruct、Llama3.1-8B-Instruct、Mistral-7B 等一系列参数量规模较小

的大语言模型作为 RCA-MCTS 的核心 LLM，评估 RCA-MCTS 推理框架在低模型参数成本场景下的故障根因推理能力。同时我们选择 Qwen2-7B 模型作为 PRM 训练的基座模型，在 3.3 所述含推理步骤标签数据集上训练 2 个 epoch。在匹配故障知识库过程中，我们使用 BGE-m3 模型评估语义相似度。在推理框架参数设置方面，RCA-MCTS 搜索树最大深度设置为 6，评测数据集中所有故障的参考运维动作序列长度均小于该值，确保框架推理表现不会受限于树深度上限；设置 MCTS 过程循环次数为 8，单次推理最大 token 输出为 2 048。上述 PRM 训练、核心 LLM 在故障根因推理数据集上的有监督微调（SFT）、推理框架的实现均在 8 张 NVIDIA H100 GPU 上进行。

### 4.3 端到端表现

我们评估了 RCA-MCTS 推理框架的端到端故障根因推理效果。如图 7（a）所示，同基于 ReAct 格式的链式推理框架相比，3 个模型均通过 RCA-MCTS 获得了根因推断准确率的大幅提升。其中，Qwen2-7B-Instruct 获得的提升幅度最大，这与其在 Baseline 上表现最好、展现的故障根因推断基础能力最高相关。同时，基于 ReAct 故障推理数据集格式的 SFT 也提升了根因推断准确率，其中 Llama3.1-8B-Instruct 通过 SFT 将故障推理能力提升至接近 RCA-MCTS 的水平。我们进一步将微调和推理框架结合，3 个模型均基于 SFT+RCA-MCTS 的范式达到最佳故障根因推理效果。

我们通过故障推理动作序列匹配度进一步评估 RCA-MCTS 推理过程的合理性。如图 7（b）所示，相比于 RCA-MCTS 推理框架，SFT 更能提升其序列匹配度。这是因为 SFT

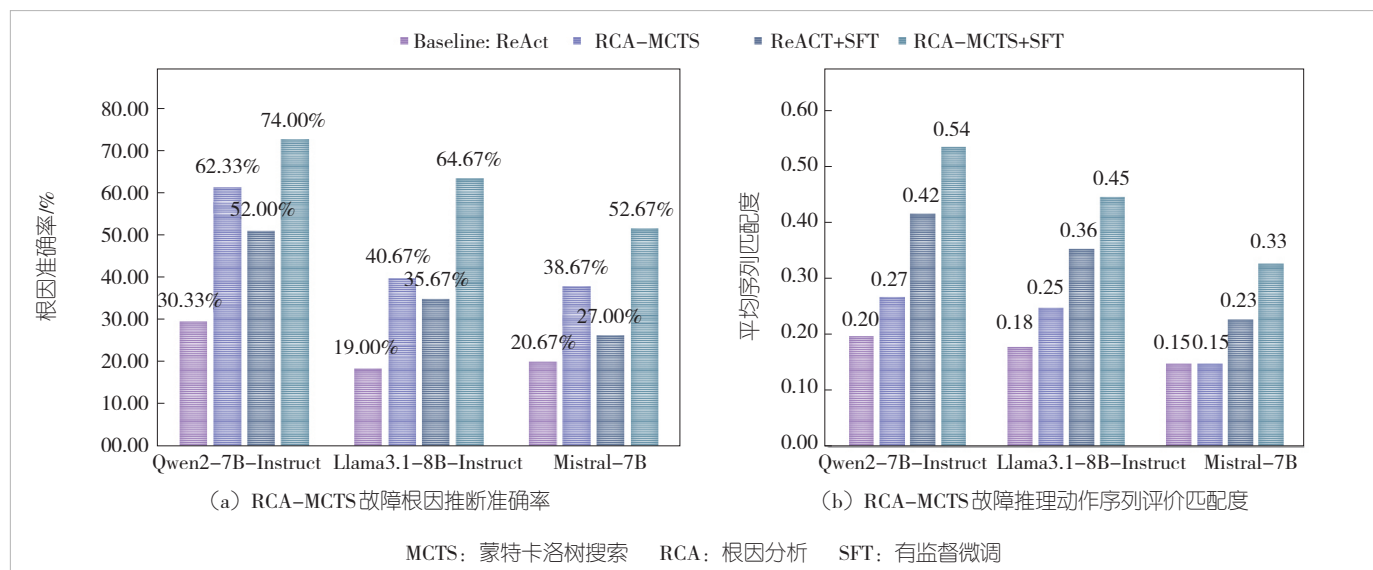


图7 RCA-MCTS推理框架端到端评测



数据集包含大量动作序列信息供LLM学习。RCA-MCTS推理框架本身不学习这些序列信息，而是有效扩展搜索空间，提高对正确故障根因覆盖可能性。在实际的故障运维场景中，对于唯一确定的故障根因可能存在多条故障运维动作序列。相较于采用何种序列，更关键的是准确找到故障根因，这更凸显了RCA-MCTS的实际应用优势。

4.4 更难评测集上的表现

同一故障告警输入可能对应于多个故障根因，由运维动作序列和故障模拟环境反馈决定。一般情况下，同一故障告警针对不同故障根因进行处理的运维动作集是相似的。尽管前述SFT训练集与评测集互不包含以确保能够客观评估框架的推理能力，但分属两个数据集的两个不同的故障根因可能对应于相同的故障告警信息。LLM通过在SFT训练集中学到运维动作知识，将有助于其在评测集中，针对同一故障告警输入但涉及另一故障根因的场景进行推理。然而，实际复杂网络系统运维场景中故障告警类别繁多，容易出现模型经验范围以外的全新故障告警，这是智能运维面临的一大挑战。为了评估RCA-MCTS在面对知识范围以外的故障告警下的性能，我们重构SFT训练集，使训练集中不包含评测集中的告警信息，更贴合实际运维场景。我们对Qwen2-7B-Instruct在不同SFT训练集上的训练后使用推理框架的表现做了对比。如表1所示，在SFT训练集的故障告警与评测集故障告警不重合时，使用RCA-MCTS推理框架的故障根因准确率低于重构SFT训练集前的最佳表现，这是评测难度提升所致。但重构SFT训练集后RCA-MCTS推理框架相对于baseline的ReAct方法仍然提升了15%的故障根因推断准确率，进一步体现了RCA-MCTS推理框架在更贴合实际运维场景的情况下对LLM故障根因推断能力的提升作用。

4.5 局限性

实际运维场景中，人工运维的推理策略根据复杂网络系统的迭代而动态调整，一些推理的策略和经验可能随着系统组件变化而与之之前的知识矛盾。要想将RCA-MCTS应用于动态变化的实际运维场景，还需要设计更复杂的扩展策略和

故障模拟环境模拟机制。另外，RCA-MCTS最终只利用了构建的故障搜索树的高反馈推理路径，对于其他推理数据缺乏应用。实际上，其他低反馈推理数据可能用于拒绝采样、策略梯度强化学习等进一步训练LLM以提高其故障根因分析能力，这有待未来研究探索。

5 结束语

通过设计适用于故障根因推理的扩展策略和模拟机制，本文得以将基于MCTS的LLM推理框架应用于RCA任务，提升了端到端故障根因推理准确率。此项成果为包括智算网络系统在内的复杂网络系统高效自动化运维做出了一定贡献，为未来自动化运维的研究探索了一种具有潜力的方向。

致谢

感谢北京中关村实验室副研究员陈力对本研究工作的大力支持！

参考文献

[1] OpenAI. OpenAI-ChatGPT [EB/OL]. [2025-02-25]. <https://chatgpt.com>

[2] JIN P X, ZHANG S L, MA M H, et al. Assess and summarize: improve outage understanding with large language models [C]//Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering. ACM, 2023: 1657-1668. DOI: 10.1145/3611643.3613891

[3] AHMED T, GHOSH S, BANSAL C, et al. Recommending root-cause and mitigation steps for cloud incidents using large language models [C]//Proceedings of IEEE/ACM 45th International Conference on Software Engineering (ICSE). IEEE, 2023: 1737-1749. DOI: 10.1109/ICSE48619.2023.00149

[4] CHEN Y F, XIE H B, MA M H, et al. Automatic root cause analysis via large language models for cloud incidents [C]//Proceedings of the Nineteenth European Conference on Computer Systems. ACM, 2024: 674-688. DOI: 10.1145/3627703.3629553

[5] WANG Z F, LIU Z C, ZHANG Y Y, et al. RAgent: cloud root cause analysis by autonomous agents with tool-augmented large language models [C]//Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. ACM, 2024: 4966-4974. DOI: 10.1145/3627673.3680016

[6] WEI J, WANG X Z, SCHUURMANS D, et al. Chain-of-thought prompting elicits reasoning in large language models [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2201.11903v6>

[7] YAO S Y, ZHAO J, YU D, et al. ReAct: synergizing reasoning and acting in language models [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2210.03629v3>

[8] SHINN N, CASSANO F, GOPINATH A, et al. Reflexion: language agents with verbal reinforcement learning [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2303.11366>

[9] HAO S B, GU Y, MA H D, et al. Reasoning with language model is planning with world model [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2305.14992v2>

[10] WU J Y, FENG M K, ZHANG S, et al. Beyond examples: high-

表1 Qwen2-7B-Instruct在不同难度SFT训练集下的故障根因推理准确率

	SFT + ReAct	SFT + RCA-MCTS
训练集中包含部分评测集告警	52%	74%
训练集中不包含评测集告警	33%	48%

MCTS: 蒙特卡洛树搜索 RCA: 根因分析 SFT: 有监督微调

- level automated reasoning paradigm in in-context learning via MCTS [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2411.18478v1>
- [11] MIN Y Q, CHEN Z P, JIANG J H, et al. Imitate, explore, and self-improve: a reproduction report on slow-thinking reasoning systems [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2412.09413v2>
- [12] ZHANG D, WU J B, LEI J D, et al. LLaMA-berry: pairwise optimization for O1-like Olympiad-level mathematical reasoning [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2410.02884v2>
- [13] ZHAO Y, YIN H F, ZENG B, et al. Marco-o1: towards open reasoning models for open-ended solutions [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2411.14405v2>
- [14] KANG J K, LI X Z, CHEN X, et al. MindStar: enhancing math reasoning in pre-trained LLMs at inference time [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2405.16265v4>
- [15] QI Z T, MA M Y, XU J H, et al. Mutual reasoning makes smaller LLMs stronger problem-solvers [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2408.06195v1>
- [16] ZHANG Y X, WU S X, YANG Y Q, et al. o1-coder: an o1 replication for coding [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2412.00154v2>
- [17] TIAN Y, PENG B L, SONG L F, et al. Toward self-improvement of LLMs via imagination, searching, and criticizing [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2404.12253v2>
- [18] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search [J]. Nature, 2016, 529: 484-489. DOI: 10.1038/nature16961
- [19] YAO S Y, ZHAO J, YU D, et al. ReAct: synergizing reasoning and acting in language models [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2210.03629v3>
- [20] HAO S B, GU Y, MA H D, et al. Reasoning with language model is planning with world model [EB/OL]. [2025-02-25]. <https://arxiv.org/abs/2305.14992v2>

## 作者简介



**罗子秋**，清华大学网络空间与网络科学研究院在读博士研究生；主要研究方向为网络系统智能运维、网络系统管理智能体开发等。



**苗宇锴**，中关村实验室助理研究员；主要研究领域为自然语言处理；2022年获SIGCOMM最佳论文奖；发表论文10余篇。



**李丹**，清华大学教授、博士生导师，教育部长江学者特聘教授，IEEE Fellow，北京高校卓越青年科学家计划项目负责人，国家“973”计划项目首席科学家，国家重点研发计划项目负责人，国家“十四五”重点研发计划“网络空间安全治理”专家副组长；主要从事计算机网络领域的研究工作；曾获教育部“青年科学奖”，以第一完成人获中国通信学会技术发明一等奖、中国电子学会技术发明一等奖；发表论文100余篇，获授权专利50余项。

# 算力智联网关键技术研究



## Key Technologies for Intelligent Computing Power Network

易昕昕/YI Xinxin<sup>1,2</sup>, 张乃晗/ZHANG Naihan<sup>1,2,3</sup>,  
刘雅承/LIU Yacheng<sup>1</sup>, 韩梦瑶/HAN Mengyao<sup>1,2,3</sup>,  
曹畅/CAO Chang<sup>1,2</sup>

(1. 中国联合网络通信集团有限公司, 中国 北京 100033;  
2. 下一代互联网宽带业务应用国家工程研究中心, 中国 北京 100048;  
3. 北京邮电大学, 中国 北京 100876)  
(1. China United Network Communications Group Corporation Limited, Beijing 100033, China;  
2. National Engineering Research Center of Next Generation Internet Broadband Service Application, Beijing 100048, China;  
3. Beijing University of Posts and Telecommunications, Beijing 100876, China)

DOI: 10.12142/ZTETJ.202502005

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250428.0945.002.html>

网络出版日期: 2025-04-28

收稿日期: 2025-03-20

**摘要:** 智算业务的快速发展, 计算资源需求急剧增长, 对网络能力提出了更高标准。聚焦算力智联网, 深入剖析数据入算、模型训练、模型下发、模型推理四大典型智算业务场景需求, 提出算力智联网关键技术要求。通过构建高效、灵活且可靠的智算基础设施, 为智算业务提供差异化承载、灵活算网调度和无损传输等核心能力。此外, 针对四大典型场景开展网络创新能力试点验证, 显著提升了人工智能模型训练效率。

**关键词:** 算力智联网; 差异化网络承载; 灵活算网调度; 无损传输

**Abstract:** The rapid development of intelligent computing services has led to a sharp increase in computing resource demands, imposing higher standards on network capabilities. This paper focuses on the intelligent computing power network and analyzes the requirements of four typical intelligent computing business scenarios: data ingestion, model training, model deployment, and model inference. It also proposes the key technical requirements for the intelligent computing power network. Building an efficient, flexible, and reliable intelligent computing infrastructure provides core capabilities such as differentiated network bearing, flexible computing-network scheduling, and lossless transmission for intelligent computing services. Moreover, pilot verification of network innovation capabilities for the four scenarios has significantly improved the training efficiency of artificial intelligence models.

**Keywords:** intelligent computing power network; differentiated network bearing; flexible computing and network scheduling; lossless transmission

**引用格式:** 易昕昕, 张乃晗, 刘雅承, 等. 算力智联网关键技术研究 [J]. 中兴通讯技术, 2025, 31(2): 31-38. DOI: 10.12142/ZTETJ.202502005

**Citation:** YI X X, ZHANG N H, LIU Y C, et al. Key technologies for intelligent computing power network [J]. ZTE technology journal, 2025, 31(2): 31-38. DOI: 10.12142/ZTETJ.202502005

### 1 智算业务发展趋势

随着信息技术的迅猛发展, 智能计算(后文简称为智算)已经成为推动全球数字化转型的核心力量之一。智算融合了人工智能(AI)、机器学习、深度学习等先进技术, 通过模拟人类的认知过程来处理复杂的数据和任务, 不仅改变了传统行业的运作模式, 也为新兴产业的发展提供了无限可能<sup>[1]</sup>。从自动驾驶到个性化医疗方案, 再到智能家居系统, 智算技术的应用无处不在, 深刻影响着人们的日常生活

和社会经济结构。

近年来, 智算业务取得了令人瞩目的发展, 并实现了关键的技术突破。据市场研究机构的相关数据显示, 全球智算市场规模预计在未来5年内, 将以两位数的年增长率持续扩张。这一增长态势的背后, 得益于云计算、大数据、物联网等基础设施建设的持续完善, 以及算法效率的显著提升。不仅如此, 网络因素在智算业务发展中的作用正日益突出。作为连接计算资源的关键纽带, 网络能够实现大规模算力资源的统一调配与高效协同。

中国高度重视算力与网络基础设施建设, 大力推动网络传输创新模式探索与数据基础设施建设进程。国家发展和改革委员会在相关文件中明确提出, 要“采用弹性带宽、任务

**基金项目:** 国家重点研发计划项目(2023YFB2904201); 中国博士后科学基金资助项目(2024M763570)

式服务、数据快递等创新模式”<sup>[2]</sup>；国家数据局也着重强调，需“有效提升东数西算网络传输效能”。

网络基础设施作为支撑智算业务的核心底座，目前仍面临着诸多技术瓶颈。具体而言，这些瓶颈主要涵盖大数据集传输过程中的效率低下问题、敏感数据集交易环节中的安全隐患，以及多样化业务场景下灵活调度与差异化承载的难题。

## 2 智算业务典型场景与需求

智算业务依托海量数据分析，结合高效智能的计算技术，为个人和企业提供解决方案。在此过程中，海量数据是基础，算法模型是核心，场景应用是价值体现。典型智算业务流程如图1所示：首先将海量数据上传至智算中心，随后利用这些数据进行大模型训练。训练过程可采用存算分离拉远模式或多智算中心协同方式，最终基于训练完成的大模型下发至边缘节点，为用户提供推理服务。

1) 数据入算。数据入算流程指将各类数据上传至算力中心，为后续训练与推理奠定数据基础。以医疗应用领域为例，医疗机构需将各科室及分院产生的海量医疗数据传输至专业医疗算力中心，医院可通过集中影像数据，借助深度学习算法辅助医生精准诊断疾病。类似地，政务办公、金融、教育等领域均存在数据入算场景——需将行业数据集中入算，以支撑综合分析及决策。数据入算场景的核心目标是确保网络具备强大传输能力（如大象流识别技术），以高效处理TB/PB级海量数据，并可依据业务需求灵活调配资源，提供弹性化、按需定制及任务式的服务模式。

2) 模型训练。训练的核心目标是通过海量数据与复杂算法，使模型习得数据中的规律和特征，进而具备特定的智

能决策或预测能力。不同训练场景适配于差异化的业务需求与资源条件，同时对网络提出多样化要求。当前训练环节主要包含3种实现方式：数据中心（DC）内训练、跨DC协同训练、存算拉远训练。其中，DC内训练场景应用最为广泛。跨DC协同训练面向多DC协同超大模型训练或碎片化算力整合出租场景，该场景需要广域无损、大象流识别与负载均衡等能力。针对数据敏感用户，训练场景可通过存算分离拉远技术，实现用户私域存储与人工智能计算中心（AIDC）之间的高效拉远训练，此场景需具备广域无损和广域数据安全保障等能力。此外，针对单AIDC算力资源不足的问题，可通过多AIDC协同训练实现算力资源整合，该场景要求网络提供无损、高吞吐、高性能算网互联能力。

3) 模型下发。模型下发是指将训练完成的模型从开发环境部署到生产环境或终端设备，以支持智能推理与决策。该过程通常涉及从智算中心向边缘推理池或企业自建数据中心等目标节点传输模型。模型下发对网络传输能力的要求与数据入算类似，需要网络具备高效的大规模数据传输能力，包括大象流识别等技术，从而确保大模型能够以任务化、高效率 and 灵活的方式传输至边缘节点。

4) 模型推理。模型推理是将训练完成的模型应用于实际业务场景中，针对新数据进行实时分析与预测，以提供决策支持或直接输出服务。模型推理需满足用户泛在接入与实时交互需求，要求网络具备广覆盖及按需定制的确性承载能力，通过应用感知、算力感知及算网一体化调度等能力，实现网络差异化、精准化承载。当前，AI推理能力已成为网络可提供的基础能力之一，且现有网络已较好支持大部分推理能力。

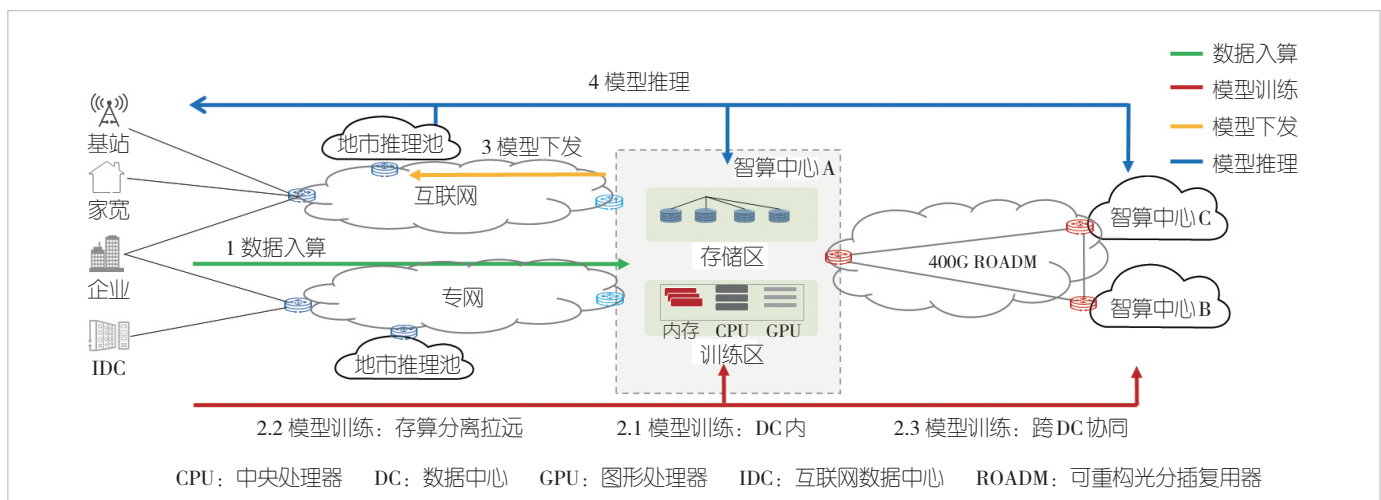


图1 智算业务典型场景



### 3 算力智联网的关键技术

智能计算业务（包括人工智能训练、大数据分析等）对网络的带宽、延迟、可靠性与安全性提出了前所未有的挑战。为支撑此类业务高效运行，网络需具备更精准的感知能力、更灵活的调度策略、更高效的传输技术及更可靠的安全保障，以实现用户与智算中心、智算中心之间的数据高效传输。通过增强上述网络能力，可构建一张具备“高通量、高性能、高智能”特征的算力智联网。

#### 3.1 应用与网络感知技术

智算业务覆盖各行各业的多样化应用，需对海量数据进行快速处理与分析。由于不同应用对网络和算力的需求各异，且复杂智能计算环境中计算任务的需求呈动态变化，因此算力智联网需具备感知能力，包括应用感知、算力感知和网络感知。

##### 1) 应用感知

应用感知的核心目标是感知应用的差异化与动态化需求，进而依据实际需求动态分配和调整资源，保障每个应用获得最优资源支持。当前，应用感知技术存在多种实现方式，如应用感知网络（APN）<sup>[3]</sup>、服务感知网络（SAN）<sup>[4]</sup>、服务化网络等。此类方案均借助互联网协议第6版（IPv6）的扩展空间增设标识，网络通过识别该标识获取应用相关信息，从而提供差异化承载服务。

应用感知网络技术包含网络侧与应用侧两种方案，如图2所示。

（1）网络侧方案。该方案由网络边缘设备对应用感知报文进行标记，再由网络设备识别并执行相应策略。该方案仅需承载网设备具备应用感知能力，无须端侧/云侧支持，部

署难度较低；但其应用感知范围局限于承载网络，且应用需求信息的精细度有限。

（2）应用侧方案。该方案由应用侧或云侧完成应用感知报文标记，要求应用客户端或云服务器支持在IPv6扩展头中嵌入应用需求信息，并由网络设备识别和执行相应策略。该方案依赖应用端和云端的生态支持，部署复杂度较高，但能实现端到端的应用感知，并更精准地将应用需求传递至网络。

智算业务的模型推理阶段呈现低时延、高带宽、低丢包率等多样化网络需求。以智能驾驶场景为例，车机与智驾模型间的数据传输具有严格的低时延要求。为此，算力智联网需精准感知应用服务的类型与需求，进而为其提供具备低时延、高带宽特性的差异化服务能力。

##### 2) 网络感知

网络感知依托时延、带宽、抖动等网络性能参数的精细化测量来实现。通过动态调整路径编排策略，网络能够显著提升数据传输效率，避免资源浪费；同时将实时网络状态反馈至应用层，支撑应用基于网络条件灵活调整传输策略。当前主流的实时网络感知方案，如随流检测（IFIT）<sup>[5]</sup>和带内操作管理和维护（IOAM）<sup>[6]</sup>等技术，与传统网络测量技术相比，通过直接测量真实数据报文，可实时、准确地呈现网络时延、丢包、抖动等关键性能指标。

IFIT作为一种带内检测技术，通过在网络真实业务报文中嵌入IFIT报文头实现对数据流的实时监测。该技术能够精准呈现网络时延、丢包、抖动等关键性能指标，实现业务故障的主动感知。同时，通过与Telemetry技术、大数据分析及软件定义网络（SDN）控制器的协同运作，可进一步构建智能运维体系。

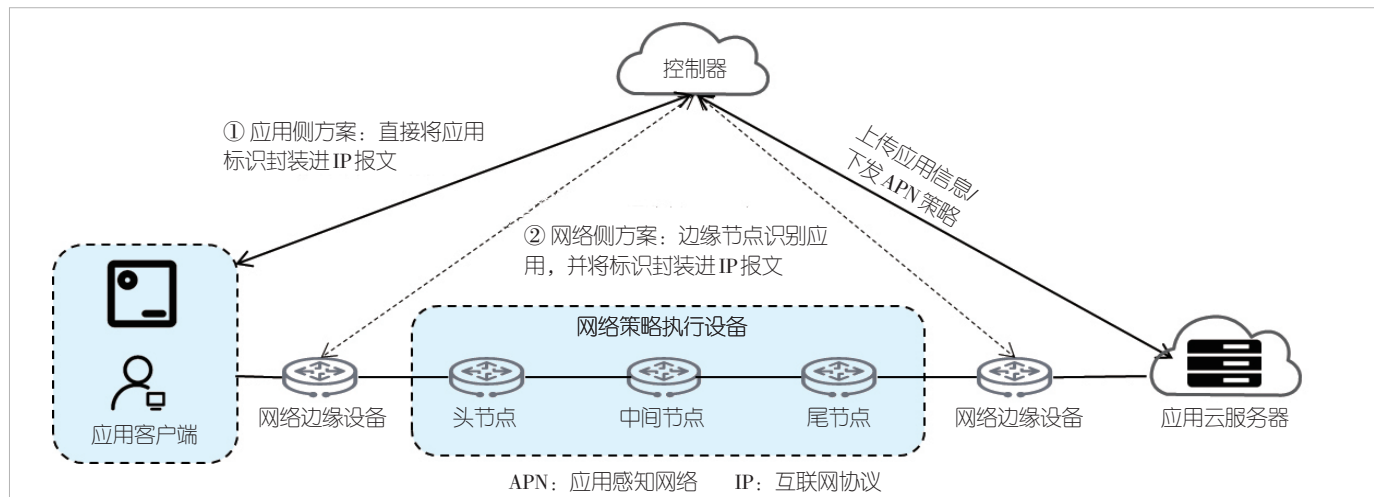


图2 应用感知网络架构

算力智联网借助增强的网络感知能力，实现网络性能指标的实时监测与智能运维管理。在精准感知应用需求后，通过精细化的网络调度策略，为智算业务提供安全可靠的网络服务保障。

### 3) 算力感知

算力感知是对智算服务节点运行状态的感知能力，旨在依据智算算力的实时状态筛选最优服务节点。在大数据与人工智能领域，计算任务通常分散于不同地理位置的数据节点。通过算力感知，能够更有效地协调这些分布式计算资源，保障数据的高效交换与协同作业。

2022年，互联网工程任务组（IETF）成立了算力感知流量调度（CATS）工作组，专门针对算力感知与路由技术展开研究<sup>[7]</sup>。当前，算力感知主要包含集中式、分布式和混合式3种方式：在集中式模式下，智算中心向上层管控层上报算力状态，由管控层统一进行路由决策；在分布式模式下，路由设备通过协议交互实时同步算力状态信息，并基于本地策略自主完成路由决策；在混合式模式下，智算中心首先将算力状态上报至管控层，再由管控层将信息下发至网络入口节点，依据业务特性，路由决策可在管控层或网络入口节点完成。然而，现阶段算力感知技术的标准化进程仍处于发展阶段，业界正积极探索并推动算力感知度量标准问题的解决。

算力智联网由用户应用、网络和智算中心构成。算力感知技术实时呈现智算中心内的算力状态，通过与应用感知、网络感知技术协同，实现端到端全链路感知，为精准捕捉智算业务应用需求、精细化承载网络流量、高效开展算网调度奠定基础。

## 3.2 跨域跨层调度技术

算力智联网需实现全网算力资源的统一整合与一体化调度。同时，为满足未来智算业务在数据传输速率、延迟、可靠性及带宽等方面不断提升的网络需求，还需具备IP与光网络协同调度的能力。

### 1) 算网一体调度

算力智联网需结合应用类型与需求，涵盖业务类别及网络性能（时延、带宽、丢包率等）要求，综合评估网络状态与智算算力资源状况，实施算网一体化调度。通过智能分析计算任务与网络路径，动态优化资源分配策略，进而提升系统整体响应速度与服务质量，为用户提供优质智算服务。

算网一体调度根据调度机制的不同，可以分为集中式调度和分布式调度。集中式调度依托中心控制节点运行。该节点负责采集全网网络感知与算力感知数据，接收应用需求信息，并基于全局视角制定资源分配和任务调度决策。此方式能够达成全局优化，保障资源利用率最大化，但中心节点的性能易成为系统瓶颈，且中心节点与用户接入点间的物理距离，可能对计算效率产生不利影响。分布式调度通过在网络各节点部署局部控制器或代理，实现分布式的资源管理与任务调度。各节点基于本地获取的业务需求、算力状态及网络状态自主决策，并与其他节点协商协作，共同完成全局任务。这种方式增强了系统的容错性与灵活性，且分布式节点通常更靠近用户接入节点，有助于提升决策效率。然而，该模式可能导致决策结果无法达到全局最优。集中式与分布式调度两种方式可根据智算业务的具体类型和需求灵活选用，亦可协同运用。

### 2) IP+光协同调度

算力智联网需依据应用感知与算力感知结果，结合业务特性，灵活选择IP网络或光网络进行数据传输。一方面，通过IP与光网络的协同调度，利用跨层优先级互感知、跨层带宽资源适配及流量调度等机制，保障端到端服务质量（QoS），提升资源利用效率；另一方面，借助精准的跨层故障定位定界与协同保护控制，实现高效的故障处理与业务恢复，有效降低运维成本。

IP+光协同调度主要涵盖管控层协同与设备层融合两大方向。在管控层协同方面，通过管控系统的联合调度与编排，实现IP路由与光路径的跨层优化。例如，基于IP层业

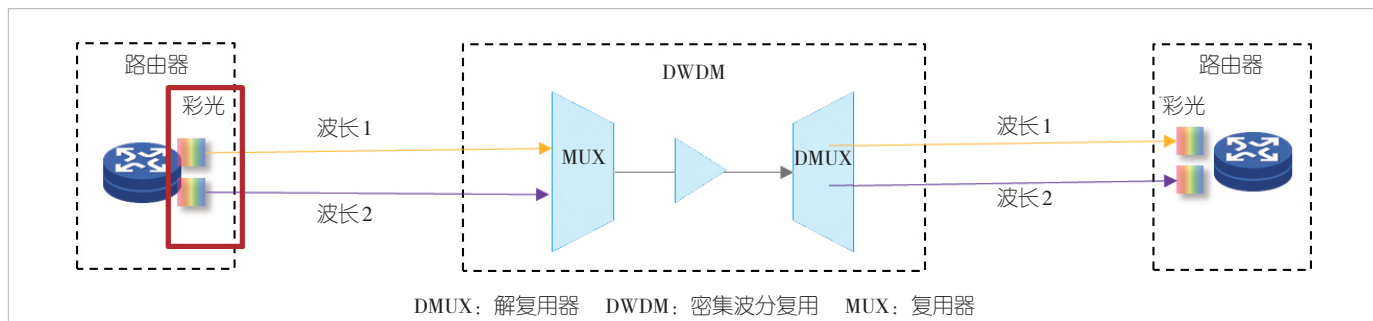


图3 路由器出彩光技术结构图

务流特征分析,动态调整光层波长路由,有效降低端到端时延。而设备层融合则依托传统路由器设备的技术创新实现跨层整合,如图3所示。传统路由器采用灰光模块连接波分设备的光波长转换单元(OTU),经OTU将灰光转换为标准彩光后,再接入复用器(MUX)进行远距离传输;与之不同,设备层融合方案将OTU模块集成至彩光模块,使路由器具备直接输出彩光的能力,可直接接入MUX设备传输,减少OTU板卡等传输设备的部署。然而,当前设备层融合技术在长距离传输场景下仍存在技术瓶颈,面临信号损耗过高、网络管控复杂等问题。目前,设备层融合的主流技术方案包括路由光网络(ROn)<sup>[8]</sup>和融合式光学路由架构(CORA)<sup>[9]</sup>。

### 3.3 数据传输技术

在智算业务中,无论是存算分离架构还是协同训练模式,均需依赖网络的零丢包传输能力。网络丢包将直接导致有效吞吐量急剧下降,从而严重影响训练效率。此外,AI训练过程中存在大象流现象(即流数量少,但单流带宽需求极大),这要求网络能够精准识别并合理拆分大象流,防止对现有网络业务产生冲击。

#### 1) 广域无损传输能力

广域网IP无损传输以降低广域承载网络丢包率、提升吞吐量、可靠性及利用率等传输能力为目标,综合运用流量控制、拥塞避免、负载均衡与网络级缓存等技术,构建基于IP承载网络的广域高通量无损传输体系。该体系可实现微秒级拥塞感知,显著降低拥塞状态下的网络丢包率,为TB级海量数据传输提供高效服务,保障应用的无损承载,对新兴的智算、超算、AI训练、分布式存储等业务发展具有重要意义。

当前,业界致力于探索和研发优化传输协议,改进拥塞控制等新技术,以实现广域无损传输。例如,基于快速用户数据报协议互联网连接(QUIC)协议的多路径传输优化方案<sup>[10]</sup>,通过动态聚合网络中多条物理链路资源,结合智能路径调度与流量分配算法,在降低传输时延的同时,大幅提升网络吞吐量与可靠性。此外,采用优先级流控制与显式拥塞通知的协同机制<sup>[11]</sup>,通过逐跳反馈抑制链路拥塞,避免全局同步引发的性能震荡,进而降低网络传输丢包率,提升数据传输效率。

#### 2) 传输协议代理能力

传输协议代理是网络通信领域的关键技术,通过在通信两端部署代理设备,实现传输协议的优化与管理,从而提升网络通信的效率、可靠性和安全性。该代理能够依据网络状态及应用需求,动态适配最优传输协议<sup>[12]</sup>。例如,在网络拥

塞场景下,代理可自动将传输协议从传输控制协议(TCP)切换至用户数据报协议(UDP)或其他专用协议以改善传输性能。此外,借助强化学习算法,传输协议代理可在多元网络环境中持续学习,优化协议选择策略。

算力智联网可借助智算中心边缘网关设备的传输协议代理功能,解决算力中心侧传输协议与网络质量不匹配的问题,通过对安全加密及传输协议进行重新封装后,在网络侧实现数据传输。具体而言,智算中心网关设备利用远程直接内存访问(RDMA)传输协议代理,对数据存储服务器发起的传输任务进行高吞吐量转发,同时能有效向端侧反馈拥塞信息,实施精细化拥塞控制,防止因网络状况恶化导致传输性能降低,进而提升网络传输效率,最大化利用网络传输带宽。

#### 3) 大象流识别与拆分

大象流(长时大流量)是指网络中持续时间长、占用带宽比例高的长连接流量,常见于视频传输、大数据同步等场景。在广域网传输中,基于传统流哈希负载分担方式,多数大象流会被分配至某一链路,极易引发网络拥塞,同时挤占老鼠流(短时小流量)的服务质量。在智算业务场景下,数据上传、协同训练等环节常出现短时大量的数据传输,若不对这些大象流进行处理,不仅会阻碍智算业务的高效传输,还会影响现网其他业务,造成拥塞、丢包等问题。因此,需对大象流进行精准识别。通过动态检测高负载流量,并实施分流或优先级调度策略,实现带宽资源的精细化分配。目前,业界主流方案是借助流量标记功能完成大象流的识别与映射:在业务头节点基于RDMA队列对(QP)进行APN ID映射,并在Segment List起始节点根据APN ID执行深度哈希处理。通过将网络中所有协议哈希因子归一化为APN ID,实现流量的识别与拆分。

### 3.4 安全增强技术

在智算业务场景下,训练数据通常存在严格的安全保密需求,如科研数据、企业核心数据等敏感信息。此类数据传输过程亟需网络具备更强的安全保障能力,可通过安全加密技术、路径溯源机制等手段,进一步强化数据传输的安全性。

#### 1) 安全加密

在传统网络加密领域,互联网安全协议(IPSec)技术常被应用于虚拟专网(VPN)构建。然而,IPSec作为点对点加密技术,仅支持基于隧道的协商机制,无法基于路由和业务实施加密,由此衍生两大问题:难以适配大型组网场景需求,无法满足基于IPv6的段路由(SRv6)加密场景要求。



为攻克这些难题，业界提出 xSEC 加密技术。该技术通过在网络边缘部署 xSEC 控制器，基于策略生成全局加密规则，实现“一次加密，全程保护”，避免多跳重复加解密操作；同时支持将加密元数据嵌入 SRv6 报文，可按段动态选择加密算法，具备更高的扩展性与灵活性。目前，业界正积极推进 xSEC 技术的标准化进程。在智算业务场景下，xSEC 技术能够为海量数据传输提供端到端安全加密服务，凭借端到端一次加密能力降低智算业务配置与运维复杂度，且可与 SRv6 技术协同实现随路加密，有效提升算力智联网的可靠性。

#### 2) 路径溯源能力

为增强算力智联网的安全性，可借助 IFIT 技术赋予智算业务路径溯源能力。该技术通过对网络实际业务流进行特征标记，实现网络性能指标的直接检测，大幅提升网络运维的及时性与有效性。IFIT 不仅能为 IP 数据包提供安全服务，实现单点加密以降低配置与运维复杂度，还可实现随路加密，显著增强网络的高可靠性。

基于 IFIT 技术，可精准定位业务的网络承载路径，实现路径溯源。该技术通过在网络真实业务报文中插入 IFIT 报文头，并结合 Telemetry 逐跳上报信息，实现对业务流端到端的实时可视化检测，进而精确还原业务流转路径。该能力对智算业务具有双重价值：可确保数据传输安全，满足智算业务的高安全性要求；通过与大数据分析及智能算法融合，构建智能运维体系，实现算力智联网的预测性分析和自愈功能，为智算业务的智能化演进提供关键技术支撑。

### 4 智算业务典型场景试验验证

依据算力智联网能力演进需求，项目

团队围绕海量数据入算、存算拉远训练、协同训练等智算业务典型场景，开展了试点验证工作。

#### 4.1 海量数据入算

面向海量数据入算场景，为增强网络调度与传输能力，项目团队在实际网络中开展了高通量数据传输技术能力验证，如图 4 所示。依托覆盖全国的骨干网络，团队选取了上海与宁夏两个节点，完成了跨越 3 000 km 的智算业务训练数据导入<sup>[13]</sup>。经测试，网络传输双向时延约为 33 ms，骨干网络的链路带宽达 100 Gbit/s，城域网带宽为 10 Gbit/s。通过应用基于 SRv6 协议的广域流量调度、传输协议优化及长距 RDMA 无损传输等技术，成功实现海量数据的任务式传输。详细测试结果与分析如下：

1) 成功实现基于 IP 承载网络的 3 000 km 海量数据任务式传输，支持用户传输带宽从 100 Mbit/s 至  $N \times 1$  Gbit/s 实现数十倍至上百倍的智能弹性调整。

2) 有效验证基于 SRv6 协议的 IP 骨干网路径可编程能力，达成上海—广州—宁夏 3 地协同、4 条并发路径的高通量数据传输。

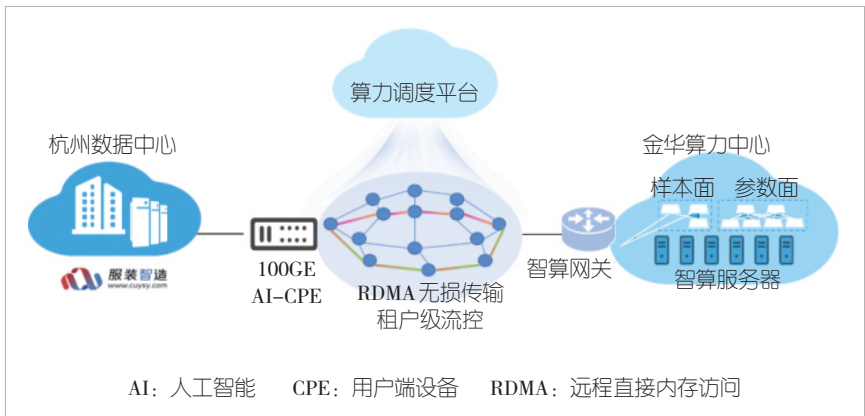


图5 存算拉远训练场景示意图

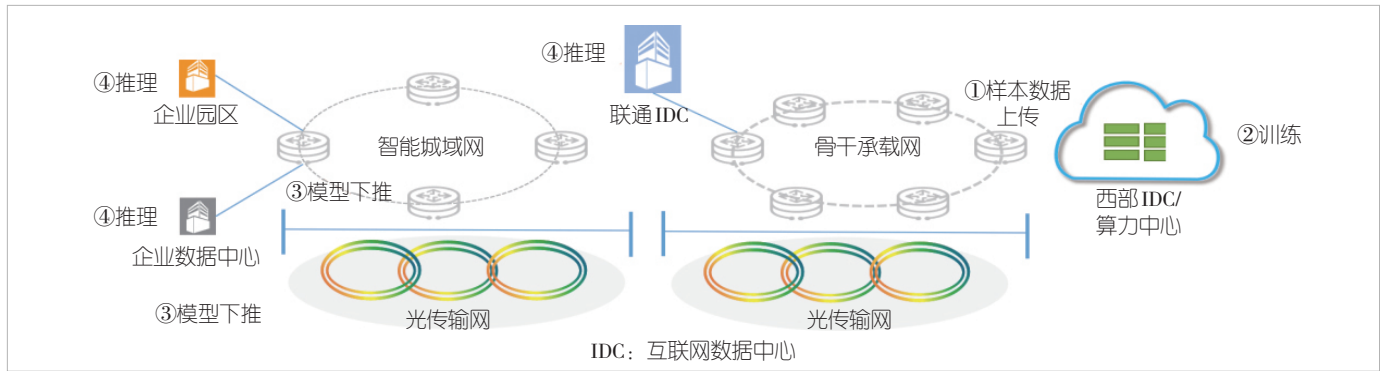


图4 海量数据入算场景示意图



3) 借助光传送网 (OTN) 无损流控与端网协同拥塞控制技术, 实现 3 000 km 长距 RDMA 流量在现网环境下的稳定传输, 端口带宽利用率从 20% 提升至 90% 以上。

#### 4.2 存算拉远训练

在智算业务场景下, 部分客户存在 AI 敏感数据本地存储、异地训练的需求。针对此类需求, 可提供存算拉远训练模式。为验证该模式的可行性, 项目团队在浙江开展了联通服装制造军团“衣瞳行业模型”的 AI 训练存算分离现网测试, 如图 5 所示。在本次测试中, 训练数据存储于杭州数据中心, 算力资源部署在金华算力中心。网络侧部署智算网关, 并运用突发流量缓存技术实现广域无损传输, 满足拉远训练过程中的数据传输要求, 达成高算力效率的存算分离拉远训练。详细测试结果及分析如下:

1) 通过运用多线程处理、增强并发性能及系统级优化等手段, 分布式存储文件系统在拉远场景下的文件读/写性能提升超 5 倍, 有效满足自然语言处理 (NLP) 模型与计算机视觉 (CV) 模型在拉远训练中的性能需求。

2) 借助长距 RDMA 广域无损传输、租户级精准流控以及存储读写性能优化等创新技术, 成功实现 30 TB 样本数据跨 200 km 的存算分离拉远训练, 计算拉远效率达 97% 以上。

#### 4.3 协同训练

在智算模型训练过程中, 当单个 DC 无法满足训练资源需求时, 需开展跨 DC 协同训练。目前, 中国联通临港智算中心已完成 AI 大模型 300 km 分布式协同训练技术验证, 充分证实了跨 DC 协同训练技术的商用可行性, 如图 6 所示。本次测试应用长距拥塞控制与精准流控协议, 在网络侧部署智算网关, 实现近端拥塞的快速识别与反馈; 同时优化智算模型并行策略, 提出适配广域带宽超大收敛比的解决方案, 有效降低超大规模智算中心互联场景下对拉远带宽的巨大需求。详细测试结果及分析如下:

1) 采用业界领先的 800G 光传送解决方案, 实现超大带宽传输与 300 km 长距覆盖, 并开展多项可靠性功能测试。验证结果显示, 智算互联需具备抗多次故障时带宽不下降的能力。

2) 借助新一代智算网关设备, 结合精准流控技术与并行方式优化技术, 达成广域收敛比不低于 16 : 1 的目标, 且

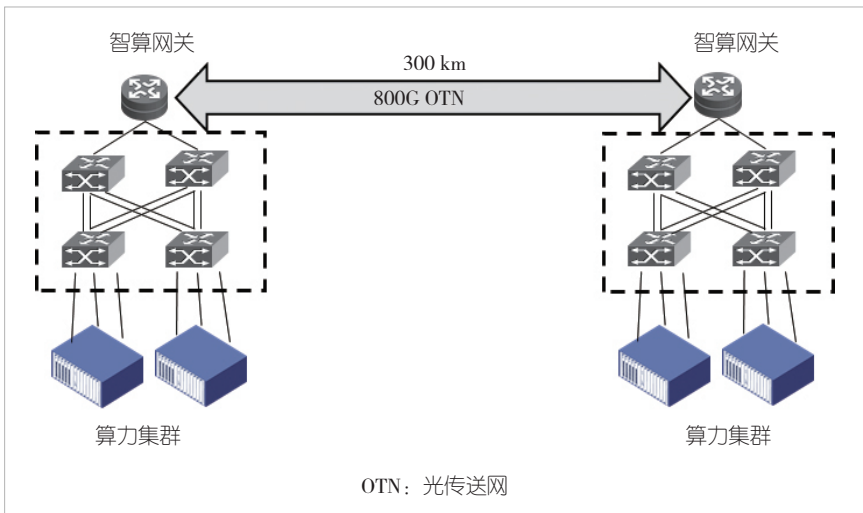


图6 协同训练场景示意图

百亿大模型分布式训练性能均达单智算中心训练性能的 95% 以上。

#### 5 结束语

针对智算业务典型场景需求, 本文中我们提出了算力智联网的新型能力演进方向, 涵盖感知、调度、传输及安全四大核心能力。通过对应用需求、网络状态、算力状态的精准感知, 满足智算业务的差异化诉求; 依托算网一体化调度与 IP+光协同调度能力的增强, 实现全网算力资源的深度整合与网络资源的高效利用; 借助无损传输能力的强化, 提升网络吞吐量, 进而加速模型训练效率。围绕海量数据入算、存算拉远训练、协同训练三大典型场景, 开展网络关键技术试点验证工作, 相关成果为未来 AI 训练模式提供了创新思路与可行方案。

#### 参考文献

- [1] 李国杰. 智能计算技术的历史性突破与巨大挑战 [J]. 集成技术, 2025, 14(1): 1-8
- [2] 国家发展改革委, 国家数据局, 中央网信办, 等. 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 [EB/OL]. [2025-03-12]. [https://www.gov.cn/zhengce/zhengceku/202401/content\\_6924596.htm](https://www.gov.cn/zhengce/zhengceku/202401/content_6924596.htm)
- [3] PENG S P, MAO J W, HU R Z, et al. Demo abstract: APN6: application-aware IPv6 networking [C]//Proceedings of IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2020: 1330-1331. DOI: 10.1109/infocomwkshps50562.2020.9162934
- [4] 中兴通讯股份有限公司. IP 网络未来演进技术白皮书 4.0——服务感知网络 (SAN) [R]. 2024
- [5] FIOCCOLA G, PANG R, WANG S, et al. Advertising in-situ flow information telemetry (IFIT) capabilities in BGP [EB/OL]. [2025-03-18]. <https://datatracker.ietf.org/doc/draft-ietf-idr-bgp-ifit-capabilities/>
- [6] SONG H, GAFINI B, BROCKNERS F. In situ operations,

- administration, and maintenance (IOAM) direct exporting [EB/OL]. [2025-03-18]. <https://www.rfc-editor.org/rfc/rfc9326>
- [7] LI C, DU Z P, BOUCADAIR M, et al. A framework for computing-aware traffic steering (CATS) [EB/OL]. [2025-03-18]. <https://datatracker.ietf.org/doc/draft-ietf-cats-framework/>
- [8] Cisco Systems, Inc. . Routed optical networking white paper [R]. 2023
- [9] Juniper. Reimaging IP over DWDM with Juniper CORA [R]. 2023
- [10] CHOUDHARY G K, KANAGARATHINAM M R, NATARAJAN H, et al. Novel MultiPipe QUIC protocols to enhance the wireless network performance [EB/OL]. [2025-03-17]. <https://ieeexplore.ieee.org/document/9120821>
- [11] LUANGSOMBOON N, LIEBEHERR J. Necessary and sufficient condition for triggering ECN before PFC in shared memory switches [J]. IEEE networking letters, 2024, 6(2): 119-123. DOI: 10.1109/LNET.2024.3382955
- [12] LIU M X, LIU Y C, MA Z F, et al. The effects of a performance enhancing proxy on TCP congestion control over a satellite network [C]//Proceedings of IEEE International Performance, Computing, and Communications Conference (IPCCC). IEEE, 2022: 325-331. DOI: 10.1109/IPCCC55026.2022.9894351
- [13] 韩梦瑶, 燕飞, 曹畅, 等. 高通量数据网演进关键技术 [J]. 中兴通讯技术, 2024, 30(6): 10-15. DOI: 10.12142/ZTETJ.202406003

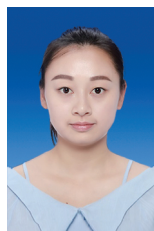
## 作者简介



**易昕昕**, 中国联合网络通信集团有限公司研究院高级研究员, 高级工程师; 主要研究领域为下一代互联网、IPv6+及智算网络等。



**张乃瞻**, 中国联合网络通信集团有限公司研究院博士后; 主要研究领域为算力网络新型架构和关键技术。



**刘雅承**, 中国联合网络通信集团有限公司高级工程师; 主要研究领域为承载网络架构与关键技术。



**韩梦瑶**, 中国联合网络通信集团有限公司研究院博士后; 主要研究领域为下一代互联网架构演进与关键技术。



**曹畅**, 中国联合网络通信集团有限公司研究院下一代互联网研究部总监, 正高级工程师; 主要研究领域为算力网络、IPv6+网络新技术、未来网络体系架构等。

# 面向多算力中心协同的广域智算网络仿真架构设计



## Wide-Area Intelligent Computing Network Simulation Architecture for Multi-Computing Centre Collaboration

边彦晖/BIAN Yanhui<sup>1</sup>, 刘明远/LIU Mingyuan<sup>2</sup>,  
虞红芳/YU Hongfang<sup>1</sup>

(1. 电子科技大学, 中国 成都 611731;

2. 北京交通大学, 中国 北京 100044)

(1. University of Electronic Science and Technology of China, Chengdu  
611731, China;

2. Beijing Jiaotong University, Beijing 100044, China)

DOI: 10.12142/ZTETJ.202502006

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.tn.20250427.1424.002.html>

网络出版日期: 2025-04-28

收稿日期: 2025-03-22

**摘要:** 针对智算仿真难以满足广域网时空动态性需求的情况, 提出了一种面向多算力中心协同的广域智算网络仿真架构。该架构的主要创新点包括: 基于属性图模型的拓扑抽象方法, 实现异构算力间不规则连接建模和不稳定网络还原; 基于流感知框架的广域通信模拟架构, 提供高精度网络通信仿真; 事件触发的多算力中心动态调度协议, 通过逻辑时钟保障跨域操作因果一致性。本架构的提出弥补了广域多算力中心背景下仿真工具的缺失, 为广域智算领域的相关研究人员提供高效、可靠的仿真支持。

**关键词:** 多算力中心协同; 广域环境; 算网融合; 仿真架构

**Abstract:** In response to the situation that intelligent computing simulation is difficult to meet the requirements of the spatio-temporal dynamics of the wide-area network, a wide-area intelligent computing network simulation architecture oriented to the collaboration of multiple computing power centre is proposed. The key innovations of this architecture comprise: 1) an attributed graph model-based topology abstraction method for modeling irregular connections among heterogeneous computing resources and restoring unstable networks; 2) a flow-aware framework-based wide-area communication simulation architecture enabling high-precision network communication emulation; 3) an event-triggered dynamic scheduling protocol for multi-computing centers that ensures cross-domain operation causal consistency through logical clocks. The proposal of this architecture makes up for the lack of simulation tools in the context of wide-area multiple computing power centers. It provides efficient and reliable simulation support for relevant researchers in the field of wide-area intelligent computing.

**Keywords:** multi-computing centre collaboration; wide-area environment; computing network convergence; simulation architecture

**引用格式:** 边彦晖, 刘明远, 虞红芳. 面向多算力中心协同的广域智算网络仿真架构设计 [J]. 中兴通讯技术, 2025, 31(2): 39-46. DOI: 10.12142/ZTETJ.202502006

**Citation:** BIAN Y H, LIU M Y, YU H F. Wide-area intelligent computing network simulation architecture for multi-computing centre collaboration [J]. ZTE technology journal, 2025, 31(2): 39-46. DOI: 10.12142/ZTETJ.202502006

近年来, 智能化浪潮席卷各行业, 人工智能 (AI) 大模型训练需求呈爆发式增长。与传统的 ResNet、视觉几何组 (VGG) 等模型相比, AI 大模型的参数量实现指数级跃升, 由此带来的算力需求也呈几何级数增长, 给计算资源带来了前所未有的挑战<sup>[1]</sup>。在自然语言处理领域, 算力需求的攀升尤为显著。以生成式预训练变换器 (GPT) 系列模型为例, 其训练过程需调用约 10 000 张 Nvidia A100 计算卡, 历经超过 60 d 方能完成训练。尽管 DeepSeek 通过创新系统

架构优化策略, 大幅降低训练所需算力, 仅使用 2 048 张 H100 计算卡即可在 54 d 内完成训练任务, 但这一量级的算力需求, 对于当前大多数已建成的算力中心而言, 依然难以满足。在此背景下, 广域网络环境下的多算力中心协同训练技术凭借独特优势, 成为业界焦点。该技术通过整合地理分布的分散算力资源, 实现数据本地化处理, 在保障训练效率的同时, 有效解决隐私保护难题。其核心在于构建广域跨算力中心的动态资源池, 突破单一算力节点的性能瓶颈; 同



时,借助智能调度机制,对异构计算资源进行全局优化配置,显著提升跨地域、跨架构计算资源的使用效率,为AI大模型训练提供更高效、安全的解决方案。

然而,广域环境下的多算力中心协同训练面临时空动态性导致的双重挑战:1)空间维度:算力资源地理分布广泛,网络拓扑呈现不规则、异构化特征;2)时间维度:跨域链路状态时变显著,导致梯度同步等关键操作的不确定性激增。

上述时空动态性特征对训练仿真工具提出了双重要求:需要精准刻画地理分布带来的拓扑异构性,又需捕获网络状态的变化规律。然而,Astra-Sim<sup>[8-9]</sup>、SimAI<sup>[10]</sup>等主流单算力中心仿真工具在该复杂场景下存在显著缺陷:1)拓扑搭建局限性:依赖规则拓扑的建模方法无法描述广域网络的不规则复合结构,导致跨域通信模拟失准;2)网络建模缺陷:现有方法将时延、带宽视为静态参数,未考虑丢包率与抖动的时空相关性;3)分布式时序混乱:基于全局时钟的假设与广域网络的异步特性相矛盾,造成分布式状态同步错位。

针对上述问题,为了还原真实广域环境下的分布式协同场景,本文提出了一种面向多算力中心协同的广域智算网络仿真架构。首先,该架构采用基于属性图模型的拓扑抽象方法与流感知框架的流量建模方法,实现对真实广域网络通信的高精度还原;其次,通过多中心间事件触发的动态调度协议,结合逻辑时钟实现跨域操作因果的一致性;最后,该架构凭借自动化模型转换框架与可视化交互界面等设计,兼具高易用性与扩展性,为用户提供端到端的便捷使用体验,为相关领域研究人员提供功能完备且操作简便的仿真工具。

## 1 面临的挑战

广域网络具有链路可保障带宽低、链路稳定性差的特点,这也是广域智算网络主要的瓶颈。与局域网相比,广域网因跨域、分布式等特点,其网络拓扑结构、性能参数等呈现高度复杂性。因此,现有针对局域网优化设计的智算仿真架构,难以在广域网环境下实现精准仿真效果,存在三大挑战:广域复杂拓扑与动态多维链路建模挑战、多中心控制面协同与同步挑战,以及系统易用性瓶颈挑战。如何应对这些挑战,以还原广域环境下的多算力中心的真实运行情况,是我们重点解决的问题。

### 1.1 广域复杂拓扑与动态多维链路建模挑战

作为智算仿真领域的开创之作,Astra-Sim将智算试验拆解为网络、计算和内存模块的设计思路,被SimAI<sup>[10]</sup>、vTrain<sup>[11]</sup>等后续工具广泛借鉴。

然而,Astra-Sim和SimAI在网络模块设计上存在显著不足,在广域场景下局限性凸显:

1)规则拓扑假设与广域复杂性存在矛盾。现有工具仅支持Ring、FullyConnected、Switch等规则拓扑的组合建模,其节点连接模式与真实广域网结构存在本质差异。这种结构性偏差在分布式训练中会导致梯度同步路径计算失准,进而造成端到端训练时间估算误差。

2)静态参数模型与广域网动态性存在矛盾。现有工具的网络通信模型采用公式(1)<sup>[9]</sup>所示的静态参数组合(LinkLatency为链路延迟,Hops为链路跳数,MessageSize为数据量,LinkBandwidth为链路带宽)。该模型未考虑时延抖动、丢包率等动态参数间的关联性,导致广域通信时间模拟存在误差。

$$\text{Time} = (\text{LinkLatency} \times \text{Hops}) + \frac{\text{MessageSize}}{\text{LinkBandwidth}} \quad (1)$$

考虑到真实广域网的特点,必须从链路拓扑抽象、链路带宽、时延、时延抖动、流量分布等多个维度对广域通信链路进行还原。因此,设计智算网络仿真架构时面临的首要挑战在于:如何实现拓扑自由搭建,以还原广域环境下节点间不规则的逻辑连接关系;同时,如何从多个维度模拟网络链路,以最大程度逼近真实网络通信场景。

### 1.2 多中心控制面协同与同步挑战

在广域网环境下的多算力中心协同计算场景中,算力中心之间同时存在数据面和控制面的通信。数据面负责各算力中心间的计算结果同步,例如在“数据并行”模式下同步前向传播的损失值,以及在“流水线并行”模式下前向传播切片结果的传递等;控制面则承担计算、通信策略与时序同步任务,以实现高效协同计算。

然而,现有的面向算力中心网的智算试验工具在设计时并未充分考虑多算力中心架构的特殊性。其通信模型仅关注数据面行为,却忽略了通信算法切换、路由策略调整等控制面操作;此外,诸如Astra-Sim等模拟器均采用全局时钟假设,该假设在广域网的高延迟环境下存在时间同步错误。因此,如何在仿真过程中模拟广域网环境下多算力中心间的控制面,体现控制面的操作逻辑并提供相应的控制行为模拟能力,成为设计多算力中心协同的仿真架构面临的第二大挑战。

### 1.3 系统易用性瓶颈挑战

考虑到仿真系统的用户需求多元化,提升系统的易用性也是设计目标之一。当前智算仿真工具存在严重的端到端流



程割裂问题。以 Astra-Sim 为例，其输入流程高度依赖第三方工具链：用户需通过 PyTorch Profile 工具和 Chakra 工具<sup>[13]</sup>预生成执行轨迹文件（Trace 模式），或借助 Scale-Sim 工具<sup>[14]</sup>重构模型代码以生成硬件参数配置文件（Txt 模式）。这种碎片化的工作流程与智算范式对灵活性的需求严重脱节。

仿真系统的易用性瓶颈还显著体现在交互设计层面。用户不仅需跨多个工具完成预处理工作，还面临可视化支持不足的难题：现有仿真系统缺乏动态链路参数配置、时序依赖关系呈现等关键交互功能，导致用户在数据预处理流程上耗费大量时间资源。因此，设计多算力中心协同的仿真架构面临的第三大挑战在于：如何构建无需第三方依赖、支持动态模型适配且符合认知效率的端到端仿真框架。

## 2 方案设计

多算力中心协同的广域智算网络呈现网络链路不稳定、网络拓扑不规则、多算力中心调度复杂的特征。为了真实还原这些特点，本仿真架构采用自下而上的设计思路，按照“拓扑还原、系统设计、应用优化”3个层次展开架构设计，具体如图1所示。

首先，基于虚拟化技术开展图属性建模和流量建模，对广域网拓扑特性和流量特性进行高保真还原，赋予仿真系统符合广域网特性的网络通信模拟能力；其次，构建“容器层、系统层、用户层”的3层架构，其中“系统层”作为中间层承上启下，向上解析用户层输入范式，向下调用容器层服务，实现算力中心间通信调度、时序同步等多种控制面功能，为仿真架构提供完整的算力仿真和网络模拟能力；最后，为提升仿真系统易用性，开发自动化模型转换框架与仿真环境自动搭建工具，满足多元化用户需求。

### 2.1 广域网络时空动态性联合建模机制

鉴于广域网络的复杂性和不稳定性，仿真架构从拓扑搭建和流量模拟两个维度构建对广域网通信的高保真还原能力。首先，基于属性图建模技术，仿真架构实现对广域环境下不规则拓扑结构的还原；其次，通过对流感知框架下的流发生器进行结构性优化，使其能够从传输层协议、带宽、数据包特征等多个维度动态生成流量，既模拟广域网可用资源

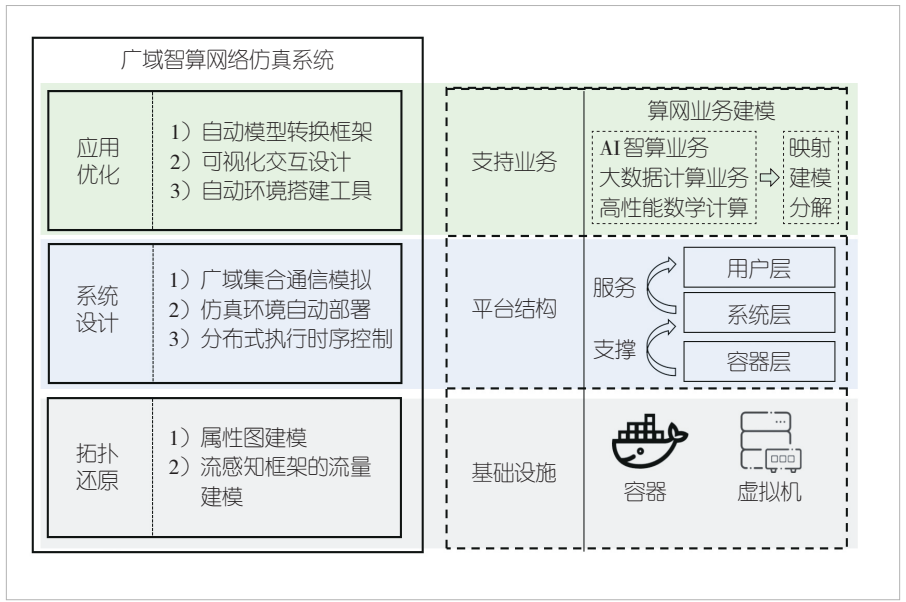


图1 广域智算网络仿真架构

的时变特性，又还原智算业务中算力中心间集合通信流量的动态变化特征。该小节公式中所涉及的参数及其含义见表1。

#### 2.1.1 属性图建模

传统仿真工具（如 Astra-Sim）因割裂处理拓扑建模与链路参数，导致广域环境下的拓扑与链路关联失真。为此，本架构提出基于属性图模型的网络抽象方法：将每个算力节点建模为图的顶点，顶点属性包含节点向量  $H = (M_{mem}, IP, N_{num})$ ，分别标识节点内存资源、IP地址、中央处理器（CPU）核数量；节点间链路建模为图边，边属性定义为参数张量  $L_i = [B_i, \tau_i, \sigma_i, \lambda_i, D_i]$ ，分别标识链路带宽、链路时延、链路时延抖动率、链路丢包率、队列大小。结合团队自研平台 Klonet<sup>[16]</sup>的自动编排技术，对抽象的属性图模型进行实体化重构，将属性图的顶点构建为容器，属性图的边

表1 属性图建模和流量建模相关的参数对照表

参数	含义	参数	含义
$M_{mem}$	内存资源	$D_i$	队列大小
$IP$	IP地址	$\mathcal{F}$	流量模型
$N_{num}$	CPU核数量	$S$	源地址
$B_i$	链路带宽	$D$	目的地址
$\tau_i$	链路时延	$P$	传输协议
$\sigma_i$	链路时延抖动	$C$	流量特征矩阵
$\lambda_i$	链路丢包率	$\mathcal{G}$	流依赖关系

CPU：中央处理器 IP：互联网协议

构建为TC-Qdisc控制的链路，其映射关系如图2所示。

通过属性图模型与流量控制（TC）规则的深度绑定，本架构实现了拓扑结构与链路参数的多维度解耦与协同控制。在拓扑构建阶段，用户可任意定义顶点连接关系，Klonet编排引擎将根据顶点属性选择合适的容器镜像，并基于边属性张量 $L_i$ 动态生成TC配置脚本。该方案保障了拓扑自由度与链路参数扩展的独立性：用户能够通过增删顶点或边灵活调整拓扑规模，彻底解决Astra-Sim等工具因硬编码拓模板导致的广域适应性缺陷；同时，每条链路凭借独立维护的五维参数 $L_i$ 进行模拟，确保仿真链路对广域真实链路的高保真还原。

### 2.1.2 流感知框架的流量建模

为实现对广域通信的真实模拟，除了设计更加符合广域网络真实情况的链路外，还需生成与之匹配的流量模式。本仿真架构针对广域网背景设计了流感知框架，并在其背景下将广域网流量抽象为五元组，即： $\mathcal{F} = \langle S, D, P, C, \mathcal{G} \rangle$ 。其中， $S/D$ 为源/目的节点的网络属性，从顶点向量 $H$ 中获取； $P$ 为传输协议且具备拓展性，支持自定义传输协议，即 $P \in \{\text{TCP}, \text{UDP}, \text{Custom}\}$ ； $C$ 为流量特征矩阵，包含包大小、包间隔、流量分布等流量特征描述， $\mathcal{G}$ 为流间依赖关系图，用于描述集合通信操作。基于该五元组描述，本仿真架构设计了适配广域网场景的专用流发生器。该流发生器采用多线程分离与实时监控技术，提高系统的并发能力，并为用户提供细粒度的流量实时监控；同时，支持包大小、包间隔和流量分布等多维空间设计，实现对广域流量的真实还原。

通过属性图建模与流感知框架的深度协同，仿真架构实现了广域网络时空动态性的高保真还原：属性图建模以顶点-边结构解耦拓扑与链路参数，灵活构建不规则拓扑并精准调控带宽、时延、抖动等五维链路参数；流感知框架则基于五元组 $\langle S, D, P, C, \mathcal{G} \rangle$ 流量模型，还原广域网流量时变性。二者通过拓扑-流量联合仿真机制，共同为仿真过程提供了对广域网背景下的通信流量的高保真还原。

## 2.2 事件驱动的多算力中心系统层调度

为了应对广域网环境下算力中心间数据同步、时序同步等多种挑战，本仿真架构自下而上设计了“容器层的系统层”“系统层的系统层”“用户层的系统层”3层结构。其中，顶层用户层接收并解析来自用户输入，将其转化为系统层规定的输入范式；底层容器层构建仿真过程的底层实体，接收并执行来自系统层的指令；而系统层作为两者之间的过渡层，在不同的任务中承担多个角色。

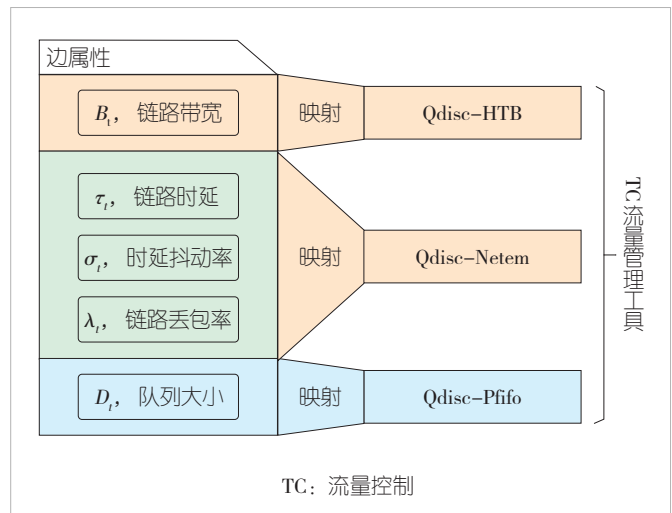


图2 边属性与流量控制工具的映射关系

- 1) 在集合通信模拟任务中，系统层将来自用户层的集合通信算法映射为点对点通信，调用容器层的通信接口进行模拟。
- 2) 针对仿真环境自动化部署任务，系统层接收来自用户层的仿真环境部署任务，将其重构为容器层搭建拓扑所需要的范式，调用容器层的拓扑接口进行模拟。
- 3) 面对分布式执行时序控制任务，系统层采用逻辑时钟对分布式任务执行进行时序控制，确保各节点执行时序的正确性。

### 2.2.1 广域集合通信操作模拟机制

系统层基于容器层构建的物理拓扑感知能力，通过动态调用容器层的接口，实现跨算力中心集合通信行为的模拟。

1) 集合通信分解映射：采用拓扑感知的通信图分解算法，将集合通信操作转化为有向无环图。图中节点表征通信任务，边表示时序依赖关系，以此确保通信仿真与真实场景的时序一致性。

2) 点对点通信建模：基于五元组流量特征，系统层调用容器层流量发生器动态生成数据流，实时采集吞吐量、时延、丢包率等指标，用于后续仿真分析。

3) 事件驱动的控制面模拟：针对集合通信操作中的控制面行为，如集合通信算法切换、路由策略切换等，系统层借助预设事件触发器，动态调整通信模式与资源分配策略，实现数据面与控制面的联合仿真。

### 2.2.2 声明式仿真环境自动部署框架

系统层提出基于拓扑语义的声明式部署框架，其核心为两个抽象阶段和一个执行阶段：

1) 拓扑抽象阶段：将用户输入的顶点-边结构映射为属性图  $\mathcal{G} = (V, E, \varphi_v, \varphi_e)$ ，其中顶点属性  $\varphi_v$  对编码算力资源（CPU/GPU 配额）进行编码，边属性  $\varphi_e$  定义多维度链路参数，涵盖带宽、时延抖动、丢包率等。

2) 资源抽象阶段：构建多维度资源联合抽象模型，通过归一化处理，将异构硬件参数映射为统一虚拟资源单元，以支持跨架构资源混合部署场景，如张量处理单元（TPU）和 GPU 集群的联合仿真。

3) 搭建执行阶段：系统层根据阶段 1 和阶段 2 的抽象结果，借助容器层提供的仿真环境部署框架，完成部署搭建。

### 2.2.3 分布式执行时序控制模型

为保障多算力中心操作的时序正确性，系统层引入分布式执行时序控制模型，其具体功能如下：

1) 基于逻辑时钟的因果排序：鉴于广域网络环境下的高延迟与强不稳定性，采用全局时钟进行多算力中心控制面调度易产生较大误差。本仿真架构基于 Lamport 逻辑时钟，构建因果排序机制，以逻辑计数代替物理时间，为系统内所有事件分配具备因果确定性的时间戳，有效解决分布式环境中因缺乏全局时钟引发的时间顺序混乱问题。

2) 乱序消息重排序处理：针对网络抖动导致的消息乱序问题，通过引入暂态因果缓冲机制，按逻辑时间戳对迟到消息重新调度，确保处理顺序严格遵循事件因果关系。

3) 关键操作的原子性保障：对于梯度同步、参数更新等关键操作，系统层通过逻辑时钟确保其原子性。通过逻辑时钟的有序递增和消息传递中的时钟更新策略，避免关键操作不受其他操作干扰，从而保证训练结果的正确性和一致性。

如图 3 所示，优化后的系统层架构通过三层协同机制，实现以下核心功能：

1) 实现广域网集合通信操作中数据面与控制面的联合仿真，模拟真实多中心调控机制。

2) 具备拓扑无关的自动化部署能力，支持分钟级千节点试验环境构建。

3) 确保多算力中心操作的强时序一致性。

### 2.3 仿真系统的高易用性设计方案

为满足仿真系统的用户的多样化需求，本系统通过自动化模型转换框架、可视化图形界面和自动试验环境搭建工具等多个设计提升仿真系统的易用性。具体而言，自动化模型转换框架实现用户输入模型向算力中心模拟所需输入格式的自动转化，为用户提供端到端的试验能力；可视化交互设计

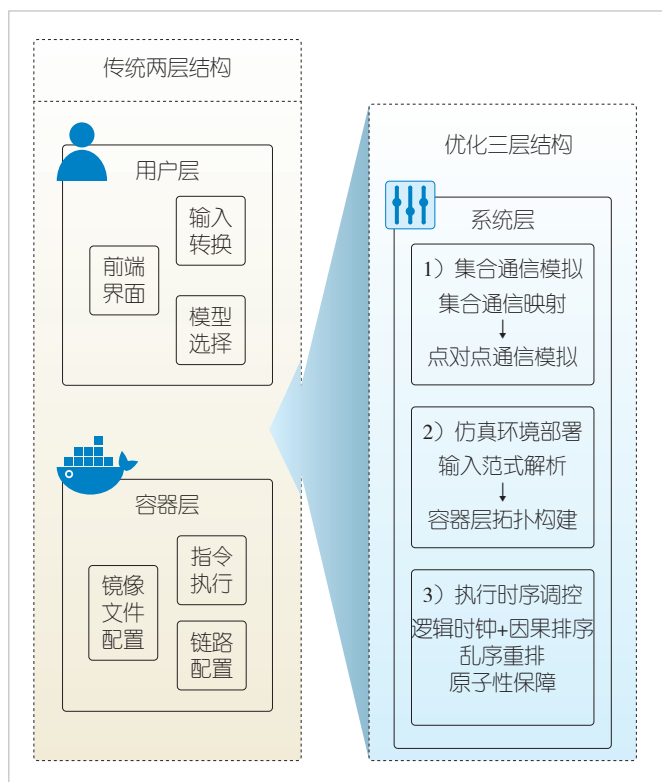


图3 优化后的三层架构

助力研究人员直观、高效地完成试验环境搭建；自动试验环境搭建工具则显著降低超大规模试验场景的搭建难度，大幅减少用户工作量。

#### 1) 自动化模型转换框架

为消除用户对第三方预处理工具的依赖，我们提出了提出分层模型转换架构。该架构支持直接解析 PyTorch、TensorFlow 等框架的模型文件，并通过内部集成的层尺寸转换器、Scale-Sim 转换器等多层转化模块，将模型文件直接输出为算力中心内部模拟训练需要的输入格式，从而为用户提供从模型文件到模拟结果的端到端全周期仿真过程。该自动化模型转换框架如图 4 所示。

#### 2) 可视化交互设计

基于团队自研的 Klonet 网络仿真系统<sup>[10]</sup>，我们设计了多算力中心智算仿真系统的可视化界面，具体如图 5 所示。用户借助拖拽、点击等交互操作，即可直观完成试验环境搭建，该方式在小规模试验场景中展现出良好适用性。

#### 3) 自动环境搭建工具

针对图形化界面难以胜任的大规模复杂试验场景搭建难题，本仿真架构配备自动搭建工具。该工具以脚本形式运行，可依据用户需求自动化构建大规模场景，有效规避了单一可视化图形界面在处理复杂大规模场景时操作繁琐的弊



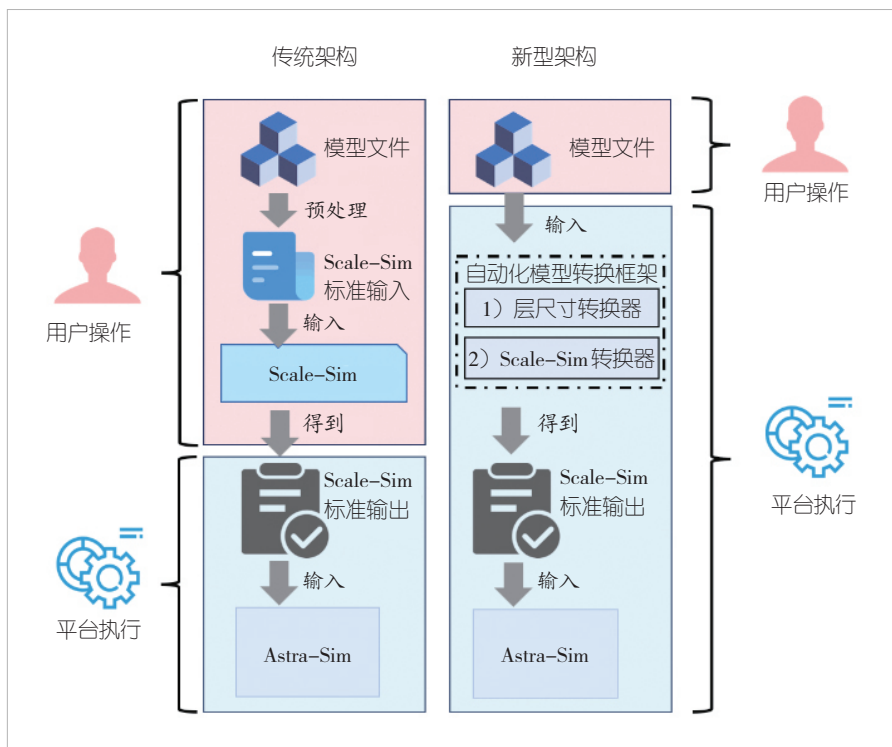


图4 自动化模型转换架构

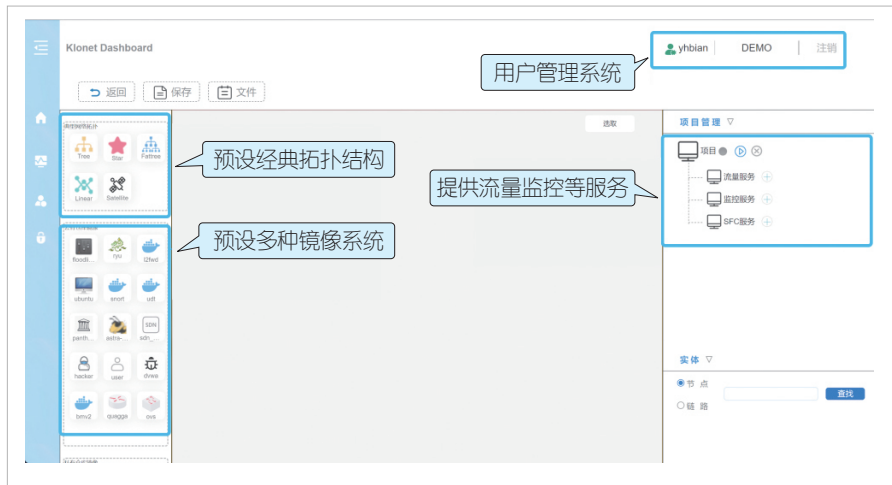


图5 可视化图形界面

端，为用户提供了多样化的场景搭建解决方案。

### 3 仿真实验

本文中，我们以广域环境下大模型多算力中心分布式训练为研究对象，开展自由拓扑搭建、试验环境配置、算力中心协同调度及广域环境多算力中心通信模拟等试验。上述试验有效验证了仿真架构在不规则拓扑构建、系统层协同调度以及广域网络高保真模拟等方面的性能。

在智算试验环节，用户可从并行方式、网络拓扑、硬件

配置、通信算法等维度进行参数设置，进而获取不同配置下的模拟训练结果。本文以 Llama-3-8B<sup>[17]</sup>大语言模型的数据并行训练为例，系统演示仿真架构的完整工作流程。

#### 3.1 输入配置

在准备阶段，用户需要准备好本次试验的配置信息，具体包括：模型文件（Llama-3-8B）、数据集（WuDao<sup>[18]</sup>）、并行方式（数据并行）、广域网络拓扑信息（此次示例中以星型结构为例）、广域集合通信算法选择（Ring-AllReduce）、算力中心内部配置信息。

用户需按照规定的范式将上述信息填入配置文件，具体如图6所示。完成配置文件后，用户可在前端界面上传该文件。

#### 3.2 环境配置

仿真架构的系统层在接收到用户层的配置信息后，对配置文件进行解析，并依次执行以下操作：

1) 确定并行模式为数据并行。系统层根据用户选择的并行模式（数据并行），调用相应的处理逻辑。例如，完成模型文件的完整分发和数据集的切分操作。

2) 搭建试验环境。系统层根据用户需求调用容器层接口，在容器层部署试验环境。具体操作包括：

- (1) 搭建不规则算力中心拓扑；
- (2) 选择算力中心所需的 Ubuntu 20.04 版本镜像和 Astra-Sim 预设镜像；

(3) 根据配置文件中的信息配置对应的链路网络信息（如带宽、延迟、丢包率等），具体如图7所示。

3) 配置文件分发。系统层将各个算力中心的配置文件分发到对应节点，并根据用户在各个算力中心的配置生成对应的执行指令。

4) 确定执行时序。系统层确定试验的整体执行时序，具体包括步骤：模型文件与数据集文件分发、单算力中心进行训练模拟、算力中心间进行梯度汇总、各算力中心进行参



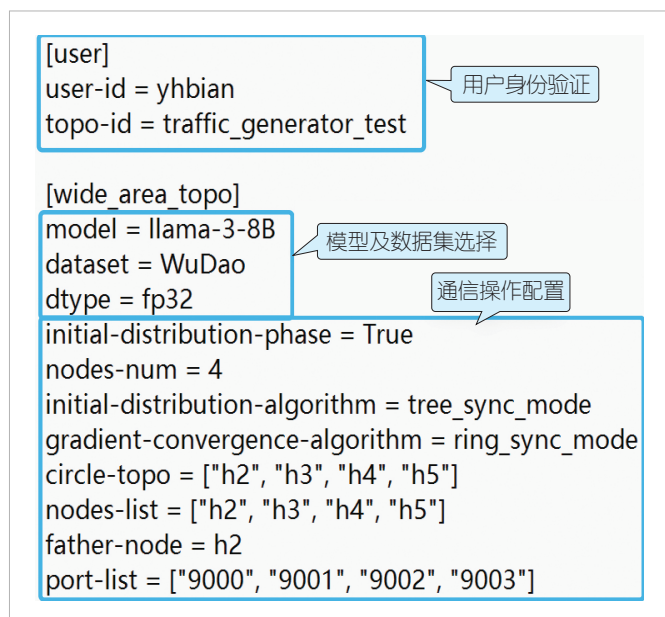


图6 用户配置文件预览

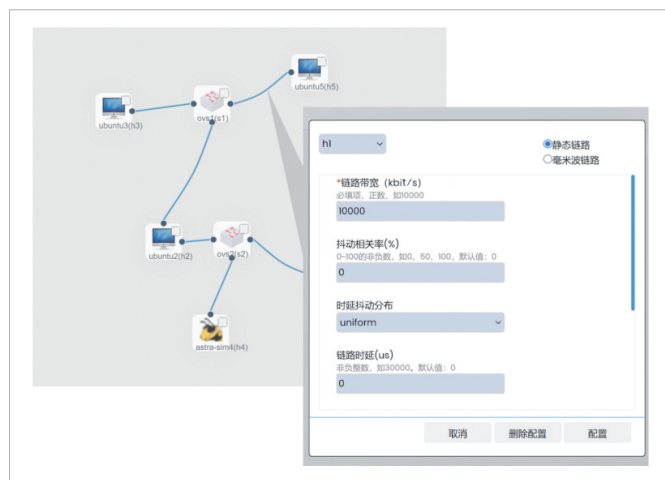


图7 搭建试验拓扑实例

数更新、进入下一个Epoch。系统层根据该执行时序依次执行对应的指令和操作。

### 3.3 仿真结果

仿真架构将模拟训练过程的结果保存至用户指定的路径,供用户后续分析和优化使用<sup>[19]</sup>。试验结果包括通信开销、计算开销、训练时间等关键指标,具体如图8所示。

1) 广播阶段的试验结果呈现了数据并行模式下, 系统在初始模型分发和训练集分发过程中的通信细节。用户可通过调整网络拓扑、变更广播通信算法等方式, 实现对广播阶段性能的优化。

2) 单个算力中心内部的模拟结果反映了各个算力中心

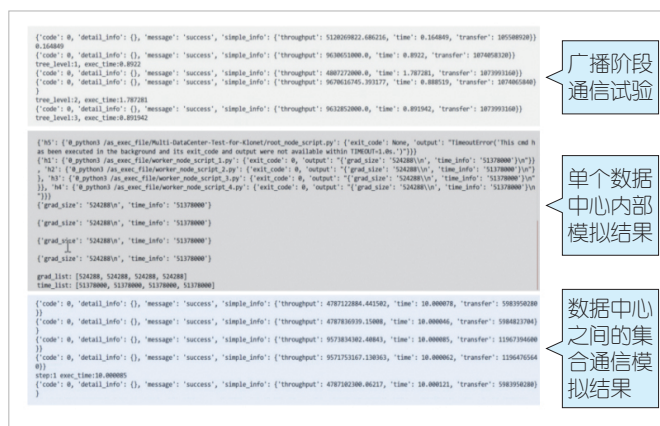


图8 仿真架构输出结果

内部的独立仿真情况。用户可以通过调整算力中心内部的算力配置、调整内部并行计算模式等方法，对算力中心的训练过程进行优化。

3) 算力中心之间的集合通信模拟结果,展示了在梯度汇聚与参数分发阶段广域环境下的集合通信操作特性。用户可以通过调整集合通信算法、优化网络配置等方法,对这一部分的通信时间进行优化,以提升整体训练效率。

## 4 结束语

本文中我们提出了一种面向多算力中心协同的广域智算网络仿真架构。该架构充分借鉴现有主流智算仿真架构的设计经验，在拓扑搭建、系统架构及易用性设计等关键层面进行创新性优化。通过构建高自由度拓扑搭建机制，设计高还原度广域网络模拟系统，并融入高易用性系统设计，为广域智算领域研究人员提供了功能完备、操作便捷的仿真平台。后续研究将持续提升架构的仿真精度与运行性能，探索新型网络技术与调度算法，并针对多样化应用场景进行适应性优化，助力推动中国AI产业发展，丰富全球技术生态。

## 参考文献

- [1] JIANG Z, LIN H B, ZHONG Y M, et al. MegaScale: scaling large language model training to more than 10, 000 GPUs [EB/OL]. [2025-03-18]. <https://arxiv.org/abs/2402.15627>
- [2] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [EB/OL]. [2025-03-19]. <https://dl.acm.org/doi/10.5555/3295222.3295349>
- [3] RADFORD A, WU J, CHILD R, et al. Language models are unsupervised multitask learners [EB/OL]. [2025-03-19]. [https://cdn.openai.com/better-language-models/language\\_models\\_are\\_unsupervised\\_multitask\\_learners.pdf](https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf)
- [4] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners [EB/OL]. (2020-07-22) [2025-03-15]. <https://arxiv.org/abs/2005.14165>
- [5] DeepSeek-AI, LIU A X, FENG B, et al. DeepSeek-V3 technical report [EB/OL]. [2025-02-18] [2025-03-18]. <https://arxiv.org/abs/>

- 2412.19437
- [6] DeepSeek-AI, LIU A X, FENG B, et al. DeepSeek-V2: a strong, economical, and efficient mixture-of-experts language Model [EB/OL]. (2024-06-19) [2025-03-18]. <https://arxiv.org/abs/2405.04434>
- [7] DeepSeek-AI, GUO D, YANG D, et al. DeepSeek-R1: incentivizing reasoning capability in LLMs via reinforcement learning [EB/OL]. (2025-01-22) [2025-03-16]. <https://arxiv.org/abs/2501.12948>
- [8] RASHIDI S, SRIDHARAN S, SRINIVASAN S, et al. ASTRA-SIM: enabling SW/HW co-design exploration for distributed DL training platforms [C]//Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS). IEEE, 2020: 81-92. DOI: 10.1109/ispass48437.2020.00018
- [9] WON W, HEO T, RASHIDI S, et al. ASTRA-sim2.0: modeling hierarchical networks and disaggregated systems for large-model training at scale [C]//Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS). IEEE, 2023: 283-294. DOI: 10.1109/ISPASS57527.2023.00035
- [10] WANG X Z, LI Q X, XU Y C, et al. SimAI: unifying architecture design and performance tuning for large-scale large language model training with scalability and precision [EB/OL]. [2025-03-18]. <https://ennanzhai.github.io/pub/nsdi25spring-simai.pdf>
- [11] BANG J, CHOI Y, KIM M, et al. vTrain: a simulation framework for evaluating cost-effective and compute-optimal large language model training [C]//Proceedings of 57th IEEE/ACM International Symposium on Microarchitecture (MICRO). IEEE, 2024: 153-167. DOI: 10.1109/MICRO61859.2024.00021
- [12] RILEY G F, HENDERSON T R. The ns-3 network simulator [M]//Modeling and Tools for Network Simulation. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010: 15-34. DOI: 10.1007/978-3-642-12331-3\_2
- [13] SRIDHARAN S, HEO T, FENG L, et al. Chakra: advancing performance benchmarking and co-design using standardized execution traces [EB/OL]. (2023-05-26) [2025-03-20]. <https://arxiv.org/abs/2305.14516>
- [14] SAMAJDAR A, ZHU Y, WHATMOUGH P N, et al. SCALE-sim: systolic CNN accelerator [EB/OL]. [2025-03-20]. <http://arxiv.org/abs/1811.02883>
- [15] HUBERT B. Linux advanced routing & traffic control [EB/OL]. [2025-03-20]. <https://www.kernel.org/doc/ols/2002/ols2002-pages-213-222.pdf>
- [16] MA T, LUO L, YU H F, et al. Klonet: an easy-to-use and scalable platform for computer networks education [EB/OL]. [2025-03-20]. <https://www.usenix.org/conference/nsdi24/presentation/ma>
- [17] DUBEY A, JAUHRI A, PANDEV A, et al. The llama 3 herd of models [EB/OL]. [2025-03-20]. <https://arxiv.org/abs/2407.21783>
- [18] YUAN A, ZHAO H, DU Z, et al. WuDaoCorpora: a super large-scale Chinese corpora for pre-training language models [EB/OL]. [2025-03-20]. <https://www.sciencedirect.com/science/article/pii/S2666651021000152>
- [19] 冯文佼, 李宗航, 虞红芳. 低资源集群中的大语言模型分布式推理技术 [J]. 中兴通讯技术, 2024, 30(2): 43-49. DOI: 10.12142/ZTETJ.202402007

## 作者简介



**边彦晖**, 电子科技大学在读本科生; 研究方向为算网融合、分布式大模型训练模拟技术。



**刘明远**, 北京交通大学电子信息工程学院讲师、信息与通信工程学院在读博士; 研究方向为网络算网融合、低空网络、多路径传输调度等。



**虞红芳**, 电子科技大学教授、博士生导师, 信息与通信工程学院副院长; 长期致力于智慧网络及应用研究; 主持多个国家自然科学基金重大项目、国家重点研发计划、军科委等国家级项目; 曾获教育部自然科学奖二等奖、电子学会科技进步二等奖、四川省教学成果二等奖等奖项, 主持研发的“跨数据中心高性能分布式机器学习系统GeoMX”和“基于轻量级虚拟化的大规模网络创新平台Klonet”分别获中国通信学会2021年未来网络领先创新科技成果奖、2021年网络5.0创新科技成果奖; 发表论文100余篇, 授权中国发明专利30余项、美国发明专利2项, 出版学术专著5本。

# 在网计算:愿景与关键技术



## In-Network Computing: Vision and Key Technologies

刘杨/LIU Yang<sup>1</sup>, 孙浩南/SUN Haonan<sup>1</sup>,  
程伟强/CHENG Weiqiang<sup>2,3</sup>

(1. 北京邮电大学, 中国 北京 100876;

2. 东南大学, 中国 南京 211189;

3. 中国移动通信有限公司研究院, 中国 北京 100032)

(1. Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Southeast University, Nanjing 211189, China);

3. China Mobile Communications Corporation Research Institute, Beijing 100032, China)

DOI: 10.12142/ZTETJ.202502007

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250425.1049.002.html>

网络出版日期: 2025-04-25

收稿日期: 2025-03-26

**摘要:** 随着5G/6G通信、边缘计算与人工智能技术的深度融合, 传统网络架构面临算力资源分布不均、服务响应延迟高、动态适应能力不足等挑战。提出三级分层在网计算体系架构, 通过算力泛在化、服务动态化与决策智能化的深度协同, 解决算网协同关键瓶颈。该架构在数据中心网络速率限制中显著提升处理性能与系统稳定性, 为高效数据传输、灵活网络服务及大规模智能训练提供支撑, 推动网络向高性能、智能化方向演进。

**关键词:** 在网计算; 架构; 算网协同; 智能网络技术

**Abstract:** With the deep integration of 5G/6G communication, edge computing, and artificial intelligence technologies, the traditional network architecture is faced with challenges such as uneven distribution of computing power resources, high service response latency, and insufficient dynamic adaptation capabilities. A three-level hierarchical in-network computing architecture is proposed. Through the deep collaboration of ubiquitous computing power, dynamic services, and intelligent decision-making, the key bottleneck of computing-network collaboration is addressed. This architecture significantly improves the processing performance and system stability in the rate limiting of data center networks, providing support for efficient data transmission, flexible network services, and large-scale intelligent training, and promoting the evolution of the network towards high performance and intelligence.

**Keywords:** in-network computing; architecture; computing-network collaboration; intelligent network technology

**引用格式:** 刘杨, 孙浩南, 程伟强. 在网计算: 愿景与关键技术 [J]. 中兴通讯技术, 2025, 31(2): 47-52. DOI: 10.12142/ZTETJ.202502007

**Citation:** LIU Y, SUN H N, CHENG W Q. In-network computing: vision and key technologies [J]. ZTE technology journal, 2025, 31(2): 47-52. DOI: 10.12142/ZTETJ.202502007

近年来, 在5G/6G通信、边缘计算和人工智能技术的协同驱动下, 网络应用场景正加速从“万物互联”向“万物智联”演进<sup>[1]</sup>。这一变革进程凸显出传统“云-边-端”三级算力架构的固有局限<sup>[2]</sup>: 虽然该架构通过分布式部署部分缓解了云端集中式计算的压力, 但其基于传输控制协议/互联网协议(TCP/IP)栈的底层设计存在本质性缺陷, 导致网络传输与计算资源之间难以实现高效协同, 最终制约了智能服务场景的性能表现与发展潜力。

当前网络架构存在以下3个核心问题:

- 1) 算网分离。现有网络仅作为数据传输的通道, 缺乏计算任务的感知能力, 这会导致算力与网络资源调度的脱节。
- 2) 资源僵化。网络设备功能硬件固化, 难以按业务需求动态调整。
- 3) 能效失衡。当前网络架构在端到端传输中存在两个关键缺陷: 缺乏计算任务感知能力, 缺少跨域资源协同调度机制。这导致数据传输路径非最优化和重复传输现象频发, 最终造成显著的能源浪费。此外, 云端集中式服务器与边缘节点的中央处理器(CPU)/图形处理器(GPU)等计算资源因任务分配不均、静态调度策略限制, 长期处于低负载状

**基金项目:** 国家重点研发计划项目(2024YFB2906602, 2024YFB2906601); 晋城市重点研发计划项目(20220101, 20230104)



态,加剧了整体系统的能效失衡。

针对上述挑战,学术界与产业界提出将在网计算作为新一代网络演进的核心范式。该范式的核心思想是将计算能力深度嵌入网络基础设施,推动网络从被动传输的“连接管道”向主动提供计算服务的“智能平面”转型。在网计算通过3个核心特征重构网络体系:1)算力泛在化,通过为网络节点赋予异构计算能力(CPU、GPU、现场可编程门阵列),实现算力资源的分布式部署;2)服务动态化,借助网络功能虚拟化(NFV)与软件定义网络(SDN)的深度融合,支持网络功能的按需编排与弹性伸缩;3)决策智能化,基于强化学习与数字孪生技术,构建网络资源的实时镜像与动态优化机制,从而更好地实现网络资源的自主调度与全局优化。

当前计算卸载与边缘智能研究虽取得一定进展,但现有方案仍存在明显不足:一是异构计算资源缺乏统一的量化标准,造成跨域调度效率低下;二是静态化的网络功能编排机制难以满足动态多变的业务需求。针对以上这些问题,本研究创新性地构建了三级分层的在网计算体系架构,有机整合分布式协同计算、智能资源调度与确定性网络保障三大技术,系统性地突破算网协同的关键技术瓶颈。本研究的创新性主要体现在以下3个方面:1)设计基于P4可编程数据平面与计算感知路由协议(CARP)的融合架构,实现了网络传输与计算处理的深度协同;2)设计了基于联邦学习的隐私保护计算框架,有效解决了跨域协同中的信任与安全的一系列问题;3)研发了基于深度强化学习的动态服务链编排算法,在典型的智能制造场景中实现了端到端时延的突破性优化。

## 1 在网计算体系架构

### 1.1 物理资源层的异构协同架构

#### 1.1.1 异构计算节点动态适配机制

物理资源层构建了由边缘服务器、智能网卡和可编程ASIC组成的三维异构计算矩阵。其中,基于x86/ARM混合架构的边缘服务器支持容器化应用部署和硬件加速模块动态扩展;配备高速外围组件互联(PCIe) Gen4接口的智能网卡实现网络协议卸载,集成可编程数据平面开发套件以提供线速数据处理能力;可编程专用集成电路实现基于RISC-V(一种开源精简指令集架构)指令集扩展架构的动态重构硬件加速单元,能够在AI推理引擎与视频编码模块等多种计算模式之间实现实时切换,可实时切换为AI推理引擎或视

频编码模块等不同计算模式。在网络架构方面,采用叶脊拓扑构建分布式计算平面:智能网卡作为边缘计算节点,脊交换机连接的边缘服务器构成区域计算中心,通过自适应拓扑发现协议实现设备动态注册与发现,确保系统具备微秒级低时延特性。

#### 1.1.2 软件定义光传输网络的弹性供给

软件定义光传输网络子系统基于灵活以太网技术实现物理链路虚拟化切片,并通过OpenConfig/YANG模型构建标准化控制平面。系统创新性地融合了可调谐光模块与弹性频谱分配算法,实现25~400 Gbit/s带宽的动态可调谐能力。控制平面部署智能资源感知引擎,实现光路损耗、色散特性和流量模式的多维实时监测,并基于优化算法完成波长路由与频谱分配的协同优化。传输层采用光子计算交换技术,通过硅基光子集成电路集成马赫-曾德尔干涉仪(MZI)矩阵,支持光域内纳秒级计算预处理和智能路由决策<sup>[3]</sup>。

### 1.2 虚拟化层的服务功能动态编排

#### 1.2.1 网络功能虚拟化的资源抽象模型

虚拟化层构建了三维资源抽象模型:在空间维度借助单根输入输出虚拟化(SR-IOV)技术穿透硬件虚拟化层,实现跨计算节点的资源池化;在时间维度运用时分复用机制实现GPU等稀缺资源的弹性分配;在功能维度构建标准化的虚拟网络功能组件库以支持灵活编排。资源调度器采用改进的装箱算法实现多维资源的最优匹配,并通过引入亲和性约束机制保障关键业务对物理资源的独占性需求<sup>[4]</sup>。

本研究归纳了硬件描述文件(HDP)元数据标准。该标准通过结构化封装方式统一表征异构硬件的关键特征参数,包括指令集架构(ISA)、加速器类型、内存带宽等硬件属性。智能服务编排系统通过解析HDP元数据,动态匹配最优硬件资源,实现从“计算意图”到“硬件能力”的精准调度与自动化部署。

#### 1.2.2 服务功能链的动态编排机制

服务功能链(SFC)编排引擎基于意图网络架构实现业务需求到服务功能图的智能映射<sup>[5]</sup>,其创新性体现在双层遗传算法设计:上层算法优化服务组件的拓扑排序,确保功能模块的依赖关系满足业务逻辑;下层算法完成物理资源的细粒度调度,结合网络演算理论,利用到达曲线和服务曲线建模保障端到端时延约束。编排系统通过增量计算原理,仅对受网络状态变化影响的服务链段进行重配置,使得动态调整

时延降至毫秒级,显著提升响应效率。

针对移动边缘场景,本研究设计了分布式服务代理机制,以满足动态多变的业务需求。在每个边缘节点部署轻量级服务网关,缓存常用功能组件并维护局部服务目录,同时支持应用层功能的快速调用(如数据预处理、实时推理等)。当中心编排器发生故障时,代理节点可基于预设的本地策略维持服务运行,并通过最终一致性机制实现状态同步,确保服务的连续性与可靠性。

此外,为扩展传统SFC的功能范围,本研究将应用层功能(如AI推理、视频处理等)纳入服务功能链编排框架。通过引入硬件描述文件元数据标准,编排系统能够动态匹配异构硬件资源,实现从计算意图到硬件能力的精准映射,支持复杂应用的高效执行。该机制突破了传统SFC仅限于L3网络功能的局限性,显著增强了服务功能链在应用层场景中的适配性与扩展能力。

### 1.3 智能控制层的认知决策体系

#### 1.3.1 数字孪生网络建模

数字孪生子系统构建了四维网络模型:物理实体层实时镜像设备运行状态,虚拟实例层精确映射服务实例的拓扑结构,规则引擎层封装智能化运维策略集,认知推理层通过数据驱动的优化算法自主实现决策优化。该模型采用混合仿真引擎,将微分方程用于连续状态的动态变化建模,并结合离散事件仿真用于捕捉网络行为的非线性特性,实现从光子层到应用层的跨层建模。

本研究创新性地融合生成对抗网络(GAN),构建动态异常流量生成器,通过持续的对抗训练机制提升数字孪生体的故障预测与异常检测能力。GAN模型生成的异常流量数据能够弥补传统数据集的不足,从而提高数字孪生体在复杂网络环境中的鲁棒性和泛化能力。

在线学习模块采用联邦学习框架,各物理节点通过本地数据训练轻量级模型,数字孪生体聚合模型参数更新全局认知<sup>[6]</sup>。这种机制既保护了节点数据隐私,又实现了全网知识共享。模型解释引擎基于注意力机制可视化决策依据,帮助运维人员理解AI模型的推理过程,提升系统可信度。

#### 1.3.2 深度Q学习资源调度算法

深度Q学习(DQL)算法设计了三重状态空间:微观层采集设备级中央处理器/内存利用率,中观层统计网络链路负载率,宏观层分析业务服务水平协议(SLA)达成率<sup>[7]</sup>。动作空间定义为多维资源分配向量,涵盖计算节点选择、频

谱分配和服务链映射策略。奖励函数引入模糊逻辑控制器,动态调整时延、能耗和资源利用率的权重系数。针对传统数据查询语言探索效率低下的问题,本研究提出一种课程学习增强机制。该机制将资源调度问题分解为3个关键阶段:设备选择、路径计算和服务部署,并为每个阶段训练专用子网络。最终通过元学习器集成各阶段策略,形成全局优化方案。在经验回放机制方面,采用优先经验重放技术,基于TD-error动态调整样本采样概率,从而有效提升算法收敛速度。

## 2 关键技术突破

### 2.1 网络计算一体化架构创新

#### 2.1.1 可编程数据平面构建

基于P4语言的数据平面编程体系通过协议无关处理架构,突破传统网络设备功能固化的技术瓶颈。该体系在数据平面抽象层构建统一的可编程流水线,支持包头解析图的动态配置机制,实现网络协议栈的运行时重构能力。同时,创新设计领域特定的中间表示,实现数据平面程序向现场可编程门阵列(FPGA)比特流、神经网络处理器(NPU)微码和ASIC逻辑电路<sup>[8]</sup>等异构硬件目标的高效编译优化。通过引入国际互联网工程任务组(IETF)标准化的In-band Network Telemetry扩展指令集,在数据包处理流程中实时同步采集链路时延、队列深度等网络状态参数,为上层网络决策系统提供即时反馈数据。

#### 2.1.2 计算感知路由协议

共用地址冗余协议突破传统路由协议的拓扑驱动范式,构建多维决策空间:计算节点负载率、任务处理时延、数据传输能耗、服务链跳数和SLA优先级权重<sup>[9]</sup>。协议核心采用改进的Dijkstra算法,通过将计算资源可用性量化为虚拟链路权重,构建了多维约束的最短路径优化模型。在路由更新机制方面,创新性地引入模糊逻辑控制,动态调节链路状态通告(LSA)的触发阈值,实现了控制平面开销与路由准确性的智能均衡优化。

### 2.2 分布式协同计算范式演进

#### 2.2.1 区块链赋能的算力市场

通过构建基于智能合约的分布式算力交易平台,并采用优化的实用拜占庭容错(PBFT)共识算法,将交易确认时延降低至秒级<sup>[10]</sup>。本研究还设计了事务链与审计链的双链架



构,其中事务链记录算力订单、合约执行等交易信息;审计链存储资源使用凭证和服务质量(QoS)证明等可验证存证数据。本研究创新性提出可验证计算承诺(VCC)协议,服务请求方通过零知识证明验证计算结果的正确性,无须获取原始数据。在算力定价机制方面,引入了期权合约机制,支持预定预留、现货竞价和弹性扩展3种交易模式,提升市场流动性。

### 2.2.2 联邦学习驱动的隐私计算

通过构建分层联邦学习架构:边缘节点执行本地模型训练,区域服务器进行模型聚合,中心节点优化全局参数。针对非独立同分布(Non-IID)数据难题,提出特征空间对齐算法,通过变分自编码器(VAE)生成跨域共享的隐式特征表示。在模型更新阶段,设计差分隐私噪声注入机制,采用自适应噪声比例调整策略,在隐私保护强度与模型精度间实现动态平衡。本研究还创新性开发模型水印技术,将设备指纹嵌入神经网络权重参数,有效溯源恶意节点。

## 2.3 智能资源调度体系构建

### 2.3.1 时空双维度预测模型

时空特征提取网络基于三维卷积长短期记忆网络(LSTM)架构构建,可在时间维度捕捉资源利用率的周期性变化规律,在空间维度建模节点间的关联性<sup>[11]</sup>。该架构引入多头注意力机制,构建时空依赖图,能够动态识别关键影响因素。针对长尾分布问题,通过设计混合密度网络(MDN)输出资源需求概率分布,提升预测鲁棒性。在在线学习模块中,集成概念漂移检测算法,当业务模式发生突变时,算法自动触发模型增量更新,保持预测精度持续稳定。

### 2.3.2 多目标优化决策引擎

基于NSGAIII算法构建Pareto最优解搜索框架,我们阐述了3层解空间映射机制:第1层采用拉丁超立方采样生成初始种群;第2层通过约束支配关系筛选可行解;第3层借助参考点引导进化方向。还创新性地引入了知识迁移策略,将历史优化解编码为概率模型,以指导新任务的种群初始化。针对高维决策变量难题,提出了变量分组进化策略,按照物理耦合关系将决策变量划分为独立进化组,从而提升算法收敛速度。

## 2.4 确定性网络保障技术

### 2.4.1 时间敏感网络增强

通过构建具备微秒级精度的全局时钟同步体系,改进广

义精确时间协议(gPTP),结合卡尔曼滤波技术对传输时延抖动进行补偿<sup>[12]</sup>。在流量调度方面,本研究创新性地融合时间感知整形器(TAS)和循环排队转发(CQF),设计混合调度策略:对关键流量采用离线规划的时间槽分配方案,对弹性流量实施在线抢占式调度。面向工业控制场景,研发帧复制与消除(FRER)机制的增强版本,支持三路径冗余传输,将丢包率降至 $10^{-9}$ 量级。

时间敏感网络(TSN)技术为在网计算提供了确定性服务质量保障的基础设施。通过精准的时钟同步和关键流量的时间槽分配机制,TSN能够确保在网计算任务的实时性和可靠性,尤其是在工业互联网、智能制造等场景中,计算任务对低延迟和高可靠性的要求极为严格。TSN的帧复制与消除机制进一步增强了计算任务的数据传输稳定性,为分布式计算节点之间的协同提供了高效传输支持。

### 2.4.2 网络演算QoS保障

通过建立随机网络演算(SNC)分析框架,将业务流建模为随机到达过程,同时将服务节点抽象为随机服务曲线<sup>[13]</sup>。在此基础上,推导出端到端时延上界的闭合表达式,并构建随机服务保障的充分条件。我们提出一种可组合的服务曲线建模方法,将复杂服务链分解拆解为基本构件,通过卷积运算聚合形成整体服务能力。此外,还开发了在线参数辨识算法,能够实时测算业务流突发参数与服务节点的剩余服务能力,进而动态调整资源预留策略。

网络演算技术为在网计算中的资源调度提供了理论支撑。通过对业务流的随机特性建模和服务节点能力的动态评估,网络演算能够实现计算任务的精确调度和资源的高效分配。特别是在多租户场景中,网络演算的在线参数辨识机制能够实时优化计算资源的分布,确保不同计算任务的服务质量需求得到满足。

## 3 典型应用场景

### 3.1 场景特征与技术挑战

在数据中心场景下,远程直接内存访问网络接口卡(RDMA NICs)对速率限制器提出了极高的要求,需精准实施拥塞控制和流量隔离策略。但商用RDMA NICs存在性能瓶颈,当启用速率限制且流数量超过1k时,其速率限制器会出现处理效率不足、调度性能下降等问题。

### 3.2 在网计算解决方案

分层速率限制架构(Tassel)的提出有效解决了上述问题。Tassel通过流级与包级两级协同控制实现了高效流量管



理。在流级控制阶段,采用基于令牌桶的全局调度算法实现粗粒度流量整形,通过自适应批处理机制动态确定各流可传输的数据包批量阈值,有效隐藏高速串行计算机扩展总线标准传输与调度排序的时延开销。在包级控制阶段,设计基于时间窗口的动态过滤算法,仅对时间戳位于 $[T_{\text{curr}}, T_{\text{curr}} + \Delta]$ 范围内的候选数据包执行精确速率检查,将包处理复杂度从 $O(N)$ 降至 $O(1)$ 。

### 3.3 技术效益与行业价值

实验数据显示,在100 Gbit/s链路环境中,Tassel可实现125 Mbit/s的稳定吞吐量。与现有方案SE-PIEO相比,Tassel的吞吐性能提升了3.6倍。当流规模扩展至16 k时,Tassel对FPGA逻辑资源与存储资源的占用分别低于5%和1%,验证了架构的硬件友好性。对比测试表明,在1 024队列对(QP)场景下,NVIDIA BlueField-3吞吐量下降63%,而Tassel保持线速传输。在时延控制方面,系统在10 k并发流时,速率控制精度仍能保持 $\pm 1.5\%$ 以内,支持100 kbit/s~100 Gbit/s的动态范围调节,满足多租户场景下SLA要求<sup>[14]</sup>。

## 4 挑战与发展趋势

### 4.1 技术挑战

当前计算资源度量体系存在维度割裂与语义鸿沟问题:传统CPU以每秒指令数(IPS)来衡量性能,GPU采用每秒万亿次浮点运算(TFLOPS)表征计算能力,智能网卡则通过每秒包处理数(PPS)评估处理效能<sup>[15]</sup>。这种离散化的度量模式引发三大核心困境:1)跨平台性能缺乏可比性,x86服务器的1 TFLOPS与ASIC加速器的同等算力无法提供等价服务能力;2)难以实现多维资源联合建模,无法有效量化计算单元与内存带宽、输入/输出吞吐之间的耦合关系;3)服务质量预测失准,单一维度指标难以准确映射至端到端时延、任务成功率等业务级SLA参数。

深层矛盾根源于冯·诺依曼架构的固有局限:计算密度、存储墙效应与能效比之间的非线性关系难以通过线性模型准确刻画。当前研究虽尝试利用等效计算周期(ECP)模型统一异构资源度量,但在处理存算一体芯片等新型架构时仍存在局限性。因此,亟需构建包含时空复杂性、能量熵和信息熵的多维度度量框架,实现“计算意图-硬件能力-服务效能”的精准映射。

### 4.2 发展趋势

量子信息技术将重构算力网络基础架构<sup>[16]</sup>,形成三级增强效应:1)安全增强,通过量子密钥分发构建具备信息论

安全的算力调度通道,有效抵御量子计算时代的中间人攻击;2)协同增强,借助量子纠缠态实现跨地域计算节点的状态同步,支撑分布式量子退火算法求解万亿变量级组合优化问题;3)效能增强,构建量子-经典混合计算架构,由量子处理器专注非确定性多项式困难(NP-hard)类问题求解,经典算力网络承担预处理与结果验证任务。

量子算力网络发展的关键突破在于量子中继技术的工程化实践:通过量子存储与纠缠交换机制实现长距离量子通信网络构建,进而打造全国性量子算力骨干网。预计2028年后,首个量子赋能的算力交易市场将应运而生,支持量子计算资源的按需租赁与协同调度服务。

## 5 结束语

本文中我们提出了在网计算体系架构通过算力泛在化、服务动态化与决策智能化的深度协同,重构了网络与计算资源的交互范式。其三级分层架构实现了异构硬件资源池化、动态服务编排及AI驱动优化的有机融合,解决了传统“云-端-边”架构的算网分离、资源僵化与能效失衡问题。在网络协议栈重构、隐私保护计算和确定性服务质量保障等关键技术领域取得创新性突破。面向未来发展,需进一步解决异构资源统一度量等基础性难题,并探索与量子通信融合的新型计算范式。该研究成果为6G网络、工业互联网等新兴领域提供了重要理论支撑与技术基础,是网络基础设施向“计算即服务”高级形态演进的重要阶段性成果。

### 参考文献

- [1] LI X, SAMAKA M, CHAN H A, et al. Network slicing for 5G: challenges and opportunities [EB/OL]. [2025-03-22]. <https://ieeexplore.ieee.org/document/8039298>
- [2] LOGHIN D, RAMAPANTULU L, TEO Y M, et al. Towards analyzing the performance of hybrid edge - cloud processing [EB/OL]. [2025-03-25]. <https://ieeexplore.ieee.org/document/8812205>
- [3] DONG W, LIU X, HE X L. Multi-stratum joint strategy based on decomposition implement in SDON [C]//Proceedings of 16th International Conference on Optical Communications and Networks (ICOON). IEEE, 2017: 1-3. DOI: 10.1109/icoon.2017.8121502
- [4] GIL HERRERA J D J, BOTERO VEGA J F. Network functions virtualization: a survey [J]. IEEE Latin America transactions, 2016, 14(2): 983 - 997. DOI: 10.1109/tla.2016.7437249
- [5] ALI ZAMANI, SHARIFIAN S. A novel approach for service function chain (SFC) mapping with multiple SFC instances in a fog - to - cloud computing system [EB/OL]. [2025-03-25]. <https://ieeexplore.ieee.org/document/8700535>
- [6] HU S G, LI C H, LI B K, et al. Digital twins enabling intelligent manufacturing: from methodology to application [J]. Intelligent and sustainable manufacturing, 2024, 1(1): 10007. DOI: 10.35534/ism.2024.10007

- [7] ZHANG C F, DONG M X, OTA K. Fine-grained management in 5G: DQL based intelligent resource allocation for network function virtualization in C-RAN [J]. IEEE transactions on cognitive communications and networking, 2020, 6(2): 428 - 435. DOI: 10.1109/tccn.2020.2982886
- [8] LI P L, LUO Y. P4GPU: accelerate packet processing of a P4 program with a CPU - GPU heterogeneous architecture [EB/OL]. [2025-03-25]. <https://ieeexplore.ieee.org/document/7856573>
- [9] YAO H J, DUAN X D, FU Y X. A computing-aware routing protocol for computing force network [EB/OL]. [2025-03-22]. <https://ieeexplore.ieee.org/document/9860158>
- [10] ZHANG Q Y, SU J, MA Z H, et al. Blockchain model testing and implementation based on improved PBFT consensus [C]// Proceedings of 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). IEEE, 2021: 1010-1025. DOI: 10.1109/idaacs53288.2021.9660959
- [11] LUO Q Y, HU S H, LI C L, et al. Resource scheduling in edge computing: a survey [J]. IEEE communications surveys & tutorials, 2021, 23(4): 2131 - 2165. DOI: 10.1109/comst.2021.3106401
- [12] SATKA Z, BARHIA D, SAUD S, et al. Experimental analysis of wireless TSN networks for real-time applications [C]// Proceedings of IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE, 2023: 1-4. DOI: 10.1109/etfa54631.2023.10275665
- [13] HU B, WANG G C. Study on delay boundary based on SNC in wireless sensor networks [C]//Proceedings of 7th International Conference on Wireless Communications, Networking and Mobile Computing. IEEE, 2011: 1-4. DOI: 10.1109/wicom.2011.6040353
- [14] WANG Z L, WAN X C, LI L Y, et al. Fast, scalable, and accurate rate limiter for RDMA NICs [C]//Proceedings of the ACM SIGCOMM 2024 Conference. ACM, 2024: 568-580. DOI: 10.1145/3651890.3672215
- [15] HONG C-H, VARGHESE B. Resource management in fog/edge computing: a survey on architectures, infrastructure, and algorithms [EB/OL]. [2025-03-26]. <https://dl.acm.org/doi/10.1145/3326066>

- [16] ARALIKATTI S. Quantum computing: challenges and opportunities [EB/OL]. [2025-03-28]. <https://ieeexplore.ieee.org/document/9616647>

## 作者简介



**刘杨**, 北京邮电大学教授, 博士生导师; 主要研究方向为算力芯片与网络关键技术; 参与多个国家自然科学基金、国家重点研发计划等项目; 联合研发了智能视觉深度学习处理器BM1684X、基于RISC-V内核的多核处理器SG2042、高性能低功耗深度学习视觉处理器CV1812H/CV1813HR等; 获2022年度山西省科技进步奖二等奖。



**孙浩南**, 北京邮电大学在读硕士研究生; 研究方向为高性能计算与网络资源协同优化、在网计算一体化架构与智能资源调度等; 参与多个国家级科研项目。



**程伟强**, 东南大学网络空间安全学院在读博士, 中国移动研究院基础网络技术研究所副所长, 教授级高工; 主要从事下一代互联网、数据中心网络以及传输网等技术研究和标准制定等工作; 参与IETF、ITU-T等10余项国际标准的制定。

# 智算网络发展综述



## Development of Intelligent Computing Networks: A Survey

段晓东/DUAN Xiaodong, 程伟强/CHENG Weiqiang,  
张昊/ZHANG Hao

(中国移动通信有限公司研究院, 中国 北京 100053)  
(The Research Institute of China Mobile, Beijing 100053, China)

DOI: 10.12142/ZTETJ.202502008

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250430.1509.002.html>

网络出版日期: 2025-04-30

收稿日期: 2025-02-25

**摘要:** 智算中心网络作为智算中心的连接底座, 需要具备高性能、低时延的通信能力。智算中心网络体系是一个多要素融合的复杂系统, 依赖于智算业务、机内/外交换芯片、网卡、网络设备等上下游产业协同创新。系统剖析了服务器/超节点内图形处理器 (GPU) 互联网络、同园区机间互联网络、跨园区智算中心互联网络三大核心领域, 探讨了智算网络的需求、挑战以及业界发展态势。中国移动创新提出全向智感互联架构 (OISA)、全调度以太网 (GSE) 技术体系、弹性以太网聚合、精细化拥塞控制、物理层安全等多项创新技术, 旨在构建超大规模、超高带宽、超低时延、超高可靠的智算中心网络, 助力人工智能 (AI) 产业发展。

**关键词:** 智算中心网络; 全向智感互联架构; 全调度以太网; 智算中心互联

**Abstract:** As the connection base of the intelligent computing center, the network of the intelligent computing center needs to have high-performance and low-latency communication capabilities. The network system of the intelligent computing center is a complex system that integrates multiple elements, relying on collaborative innovation among upstream and downstream industries such as intelligent computing services, forwarding chips, network cards, and network equipment. This paper systematically analyzes three core areas, namely the graphics processing unit (GPU) Internet in servers/super nodes, the Internet between computers in the same park, and the Internet between intelligent computing centers across parks. The requirements, challenges, and development trends in the industry of intelligent computing networks are discussed. China Mobile has innovatively proposed technical systems such as omni-directional intelligent sensing express architecture (OISA) and global scheduling ethernet (GSE), as well as a number of innovative technologies including elastic Ethernet aggregation, refined congestion control, and physical layer security. The goal is to build an intelligent computing center network with super scale, ultra-high bandwidth, ultra-low latency, and ultra-high reliability, so as to boost the development of the artificial intelligence (AI) industry.

**Keywords:** intelligent computing center network; omni-directional intelligent sensing express architecture; global scheduling ethernet; intelligent computing center interconnection

**引用格式:** 段晓东, 程伟强, 张昊. 智算网络发展综述 [J]. 中兴通讯技术, 2025, 31(2): 53-62. DOI: 10.12142/ZTETJ.202502008

**Citation:** DUAN X D, CHENG W Q, ZHANG H. Development of intelligent computing networks: a survey [J]. ZTE technology journal, 2025, 31(2): 53-62. DOI: 10.12142/ZTETJ.202502008

## 1 AI 业务与智算中心网络的发展

### 1.1 AI 业务发展趋势

在当今数据驱动的时代, 人工智能 (AI) 大模型正掀起一场前所未有的网络革命, 成为推动智能时代前行的新浪潮。随着 AI 技术的飞速进步, 大语言模型 (LLM) 以其前所未有的能力已迅速成为行业焦点。

大模型遵循三大统计特征: 第一, Scaling Law<sup>[1]</sup>, 即模型表现依赖于模型规模、计算量和数据量, 这些因素之间呈

现幂律关系; 第二, Chinchilla Scaling Law, 模型大小和数据量要同等比例扩展, 即数据量需达到参数量的 20 倍, 模型训练结果才能达到饱和; 第三, 涌现能力<sup>[2]</sup>, 产业界发现, 大语言模型相比之前的预训练语言模型 (PLM) 最显著的特征之一是它们的涌现能力。随着模型规模的指数级增长, 模型内部出现了在小规模模型中难以观察到的涌现能力。这种涌现能力使得 AI 执行任务的能力大幅提升, 且该提升与具体模型无关<sup>[3]</sup>。这种由量变引起的质变, 不仅体现在处理复杂任务的能力上, 更深刻地改变了我们对智能计算中心 (智算中心) 网络设计和优化的理解。为了不断提高模型的表现能力, AI 大模型发展迅速, 参数量不断增加。以 Chat GPT 为例, 据统计<sup>[4]</sup>, GPT-1 由 1.17 亿个参数组

基金项目: 国家重点研发计划项目 (2024YFB2906600)



成, GPT-2 中的参数数量增长至 15 亿, GPT-3 中的参数数量为 1 750 亿, 而 GPT-4 则约有 1.8 万亿个参数, 其参数数量约是 GPT-3 的 10 倍, 是 GPT-1 的 15 000 倍。

由于存在算力、内存的限制, 为了提高训练效率和模型泛化能力, 大模型需要通过多张图形处理器 (GPU) 卡进行并行训练。常见的并行训练模式有数据并行 (DP) 和模型并行 (MP), 其中模型并行又分为张量并行 (TP)、序列并行 (SP)、混合专家 (MOE) 并行和流水线并行 (PP)。

1) TP 对模型层内做划分, 将大型矩阵运算 (即张量操作) 分割成更小的部分, 在不同的 GPU 卡上并行执行, 以解决由模型数据导致的内存瓶颈问题。

2) SP 用于解决非模型数据 (如中间特征值) 导致的性能瓶颈问题, 将长序列训练任务分解成多个子序列块并将其保留在不同 GPU 卡中, 使其能够处理更长的输入数据。

3) MOE 并行将大模型拆分成多个小模型 (即专家), 每轮迭代根据输入样本通过门控网络, 分配给一部分专家进行计算, 以提高模型训练效率。

4) PP 将模型划分为多个层, 把不同的层按顺序分配到不同的节点上, 在拆分模型的边界处插入通信步骤, 并在层之间点对点传递中间结果 (包含计算结果的前向传递和梯度的反向传递)。PP 对带宽要求低, 对时延要求不太高。

5) DP 将大规模的数据集划分为多个子集并按批次分配给不同节点。每个节点都独立对各自的子集进行相同的计算

操作, 然后在各个节点之间汇总得到最终结果。GPU 卡间需传输大量梯度数据, 对带宽要求较高。由于通信可以被计算掩盖, 因此 DP 对于时延的要求相对较低。

在实际训练中, 可结合 DP、MP 的优势, 形成混合并行策略。例如, 除了在单机内使用 TP 和 SP 组合的策略外, 还可使用 PP 策略实现跨多台机器计算, 最后通过 DP 增加并发数量。

## 1.2 智算网络发展趋势

大模型的快速迭代演进和算力资源需求的快速增长, 驱动着智算中心组网规模向万卡级甚至 10 万卡级演进。支撑这类场景的网络通常属于高性能网络范畴, 可同时满足超 10 万节点的组网规模、太比特每秒以上的有效带宽、零丢包、亚微秒级时延和抖动、超高稳定性等极致性能要求。如此大规模的 GPU 互联互通网络共分为 3 个部分: 同服务器/超节点内的 GPU 卡间互联网络、同园区机间互联网络、跨园区智算中心间互联网络, 如图 1 所示。这 3 层网络在延迟、带宽、可靠性方面都存在较大差异。大模型根据这 3 层网络的能力进行模型切分并完成分层交互。每一层网络都有不同的性能要求和挑战, 需要分别进行优化。

1) 同服务器/超节点内的 GPU 卡间互联网络主要解决单服务器内部或超节点内部 GPU 之间互联互通问题, 承载 TP、SP、MOE 并行参数同步等紧密协作任务。TP、SP、MOE 并行时, GPU 卡间因需同步大量参数数据, 且通信和计算不可以掩盖, 对互联网络有纳秒级超低延迟、太比特每秒级超高

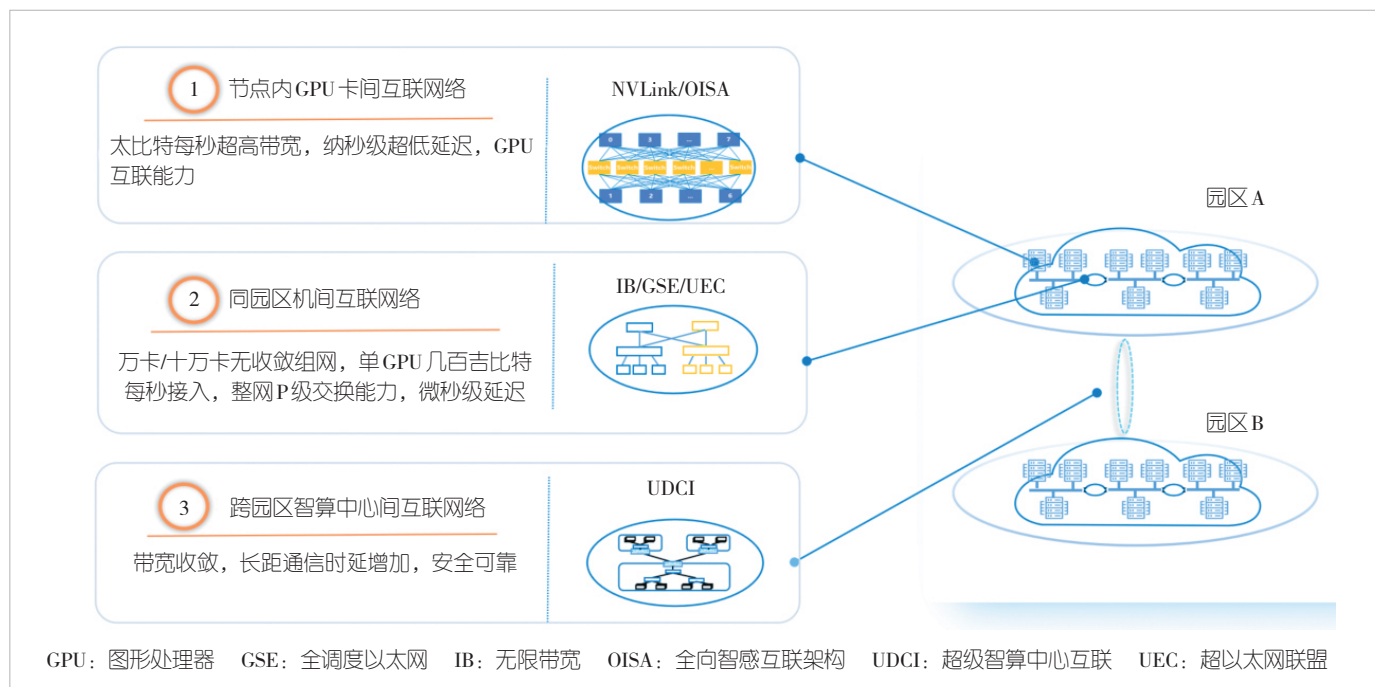


图1 新型智算中心网络

带宽和强一致性需求。机内互联网络的性能直接影响TP、SP、MoE的并行效率，其硬件实现通常依赖专用高速接口，如NVLink或外围组件互连高速总线（PCIe）通道。当前中国智算服务器只支持8卡PCIe直连拓扑、点到点带宽小于64 GB/s，无法满足MoE并行高频率、大带宽、低延时的通信需求。为此，产业界正在积极开展超节点服务器的研究。

2) 同园区机间互联网络主要解决同一智算中心内的跨服务器/机架通信问题，承载PP、DP并行参数或梯度同步任务。一个千亿级大模型训练需要万卡规模的GPU协同、单GPU几百Gbit/s的高带宽接入，并通过远程直接内存访问（RDMA）来减少传输时延。机间互联网络需要具备可规模扩展并支持万卡甚至十万卡无收敛组网。整个智算中心网络支持P级别交换能力和微秒级低延迟，通常使用IB和基于融合以太网承载远程直接内存访问（RoCE）网络技术。IB在市场中占据重要地位，但其产业开放性、部署成本面临挑战；而RoCE底层采用传统以太网，产业开放性好，但性能受到制约，基于传统以太网流式转发的优化，性能提升有限，且不适用于稀疏模型。随着AI大模型的快速发展，GPU集群规模从千卡向万卡、10万卡演进，稀疏模型也得到广泛应用。智算中心网络技术成为智算中心的瓶颈，需要从基本的转发机制进行革新，主要通过优化负载均衡和主动拥塞控制来实现网络高吞吐无损传输。

3) 跨园区智算中心间互联网络主要解决跨数据中心的服务器间通信问题。因基础设施条件受限，同园区资源无法支撑更大规模GPU卡的建设。借助已建智算中心的利旧方式，多个跨地域智算中心通过广域网互联共同支撑更大的集群训练。跨园区智算中心间互联网络主要承载跨智算中心服务器/机架间PP、DP数据，通过广域网互联，使用路由器与高速传输网络设备承载。因广域网存在带宽收敛、长距通信时延（5  $\mu$ s/km）以及安全隐患等问题，网络传输需通过优化负载均衡、拥塞控制及数据加密技术，实现安全、高吞吐的传输效果。同时，训练平台应采用分层调度优化等方式，减少跨数据中心的通信数据量，并通过计算尽量降低通信时延，将长距离传输的负面影响降至最低。目前，业界正积极探讨长距跨智算中心训练的可行性。

## 2 服务器/超节点内GPU互联网络

### 2.1 服务器/超节点内GPU互联网络的需求与挑战

随着人工智能技术的迅猛发展，超万亿参数的大模型逐渐成为颠覆型技术，在处理诸如语言理解、视频生成、科学计算等复杂任务时展现出超凡能力。然而，这种能力的发挥对计算架构提出了前所未有的挑战。万亿大模型的典型架构

是MoE结构。MoE通过将大模型分解为多个“专家”子网络，并按需动态调用这些专家，来达到以少量的算力资源获得更高模型能力的效果。尽管MoE在理论上能够显著提升模型的参数量并扩大计算规模，但在实际部署时仍依赖底层硬件提供强大的“专家”交互通道。当前，智算服务器只支持8卡直连拓扑（如图2所示），无法满足MoE模型高频率、大带宽、低延时的通信需求。为此，产业界正在积极开展超节点服务器的研究。这类服务器的核心，在于其能够利用服务器内的高带宽域（HBD），构建GPU间高效的Scale up互联，有效提升张量并行、MoE并行和流水线并行的传输效率，在不改变GPU芯片物理条件的前提下实现整体性能的优化。GPU Scale up互联的通信效率，尤其是带宽和延迟的优化，已成为业界的热点方向。

### 2.2 服务器/超节点内GPU互联网络技术业界发展情况

目前，GPU互联协议主要包括PCIe、Nvidia NVLink、AMD Infinity Fabric与UALink等。其中，PCIe由Intel公司于2001年提出，主要用于连接中央处理器（CPU）与硬盘、网卡等设备，是当前应用最广泛的外设互联协议。2022年1月，PCIe 6.0正式发布，其单通道速率达到64 GT/s，x16通道下双向传输速率达到256 GT/s。PCIe 6.0放弃了前代PCIe使用的不归零编码（NRZ）调制方式，采用四电平脉冲幅度调制（PAM4），为应对PAM4高误码率挑战引入了前向纠错（FEC）机制和循环冗余校验（CRC），以进一步提升数据传输的可靠性。PCIe可以支持多个GPU通过CPU或者PCIe Switch实现连接，但这种设计并不能满足日益提升的GPU互联需求：一是，GPU间无法点对点（P2P）通信，时延较长；二是，由于CPU提供的PCIe通道数量有限，当网卡等设备对PCIe通道的需求增加时，将挤占GPU占用的通道数，导致互联的GPU数量受限。此外，由于PCIe需要兼容多种低速设备，其带宽提升速度无法满足GPU通信需求，因此业界很少依赖PCIe搭建GPU间的高速互联通道。

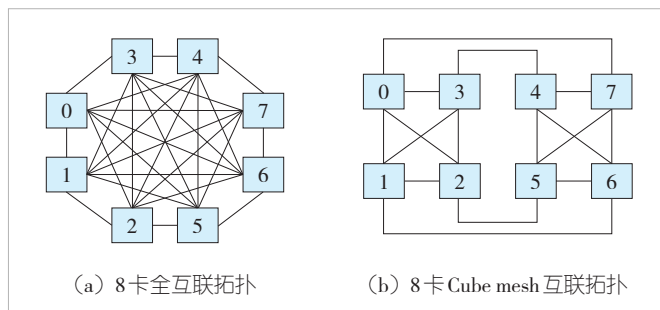


图2 传统8卡直连拓扑结构

支持GPU直接通信的互联协议可分为私有协议和开放协议两种。私有协议包括Nvidia NVLink和AMD Infinity Fabric。其中，NVLink是Nvidia在2016年P100产品中首次引入的。NVLink 1.0由成对的Sub-Link构成，每个Sub-Link包含8个差分信号对，以20 Gbit/s的速度传输NRZ的差分电信号，实现单链路40 Gbit/s的双向带宽。P100支持4条链路，总双向带宽达到160 Gbit/s。NVLink 1.0同时支持CPU-GPU以及GPU-GPU的P2P通信，允许GPU直接读写远端CPU的主机内存和相连GPU的设备内存，为高性能计算和AI应用提供了更强的数据传输能力。2024年3月，NVLink演进到了第5代，并用于Nvidia GB200 NVL72超节点服务器。从架构上看，NVLink 5.0通过NVSwitch交换芯片实现了72个Blackwell GPU和36个Grace CPU的全互联，如图3所示，其中每个GPU的双向吞吐量达到1.8 TB/s，是PCIe 5.0带宽的14倍以上。NVLink和NVSwitch为GPU集群带来了更高的通信带宽和更低的延迟，从而提升了系统的整体性能和效率。

AMD的Infinity Fabric在设计之初主要用于实现CPU与CPU的高速互联。Infinity Fabric支持每条链路42 Gbit/s的双向带宽，融合了数据传输与控制功能，由可扩展数据网络(SDF)和可扩展控制网络(SCF)两个独立通信平面构成。其中，SDF负责计算核心、内存和I/O间的数据传输一致性，SCF提供系统配置和管理的通用命令与控制机制。2021年，Infinity Fabric升级到3.0版本，并实现了CPU与GPU间的一致性互联，支持CPU的DRAM内存与GPU的HBM内存的一致性内存架构。此外，不同于英伟达NVLink仅限于内部使用，AMD已经逐步向其合作伙伴开放Infinity Fabric系统，以完善技术生态布局。

2024年由AMD牵头主导，亚马逊、Astera Labs、思科、谷歌、惠普、Intel、Meta和微软等行业巨头参与成立UAL-

ink联盟，旨在开发一种新的开放行业标准，专注于开发GPU卡间互联系统，打破英伟达垄断。UALink初版协议来自AMD，最多支持1 024个GPU互联。此外，UALink定义了创新的I/O架构，具备高性能内存语义访问原生支持，可实现显存共享、支持Switch组网模式、具备超高带宽和超低时延能力，在性能和GPU互连规模上比肩Nvidia NVLink。UALink 1.0正式协议在2025年4月发布。

### 2.3 OISA核心技术与创新点

相比于Nvidia基于NvLink与NvSwitch提供256张GPU的高速互联能力，国产相关技术目前仅支持8卡直连拓扑，交换技术的缺失导致GPU无法向超节点演进。针对当前国产GPU卡间互联带宽低、时延高、互连规模受限等问题，中国移动原创提出全向智感互联OISA协议体系，研究GPU间的Scale Up互联技术，定义物理层、数据层、事务层标准，实现超节点内多GPU芯片的对等全互联。

OISA具有统一报文格式、多语义融合、多层次流控重传以及支持集合通信加速等关键技术特征，OISA协议栈及各层功能如图4所示。GPU片上互联可基于消息语义或内存语义实现。OISA同时设计了两套报文格式，能够根据具体的应用场景选择最合适的通信模式，从而优化性能和资源利用率。在报文可靠性传输方面，OISA引入事务层选择性重传和数据层重传两种机制。其中，事务层选择性重传功能基于滑动窗口实现，能够精确识别并仅重传在互联过程中丢失或损坏的数据包，而不是无差别地重传整个数据流。与传统Go Back to N重传机制相比，OISA通过选择性重传机制提高了传输效率并减少了资源浪费，有助于实现高效物理传输设计和系统扩展性的目标。数据层重传技术能够在数据层检测到错误，并立即触发表点对点的重传机制，避免错误影响整个协议栈，从而实现数据传输的低时延和高可靠性。在流量控制方面，OISA使用了流量感知机制，支持基于信用的流量控制(CBFC)和

基于优先级的流量控制(PFC)两种流量控制机制，通过实时监控和流量模式分析，动态调整资源分配和数据传输速率，确保数据流的最优分配和信息提取。CBFC和PFC分别基于信用额度和数据优先级来精确控制数据流，有效减少网络拥塞和丢包的可能性，提高数据传输的效率和可靠性。最后，OISA支持集合通信加速技术，并将部分集合通信加速能力卸载至交换芯片侧，减少GPU与Switch的频繁数据搬运，降低模型训练和推理的时延。

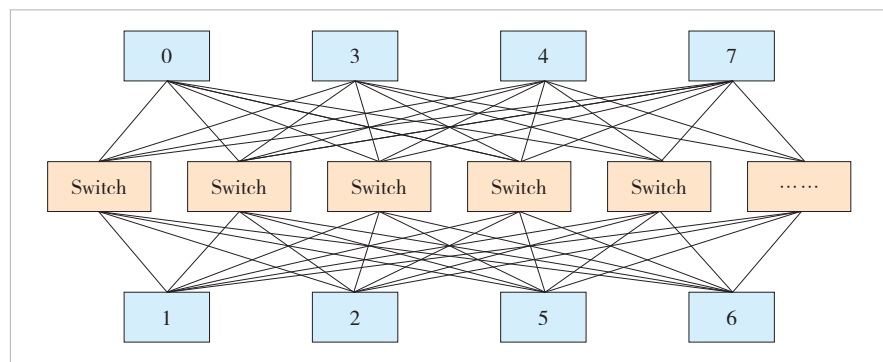


图3 基于Switch的交换拓扑结构



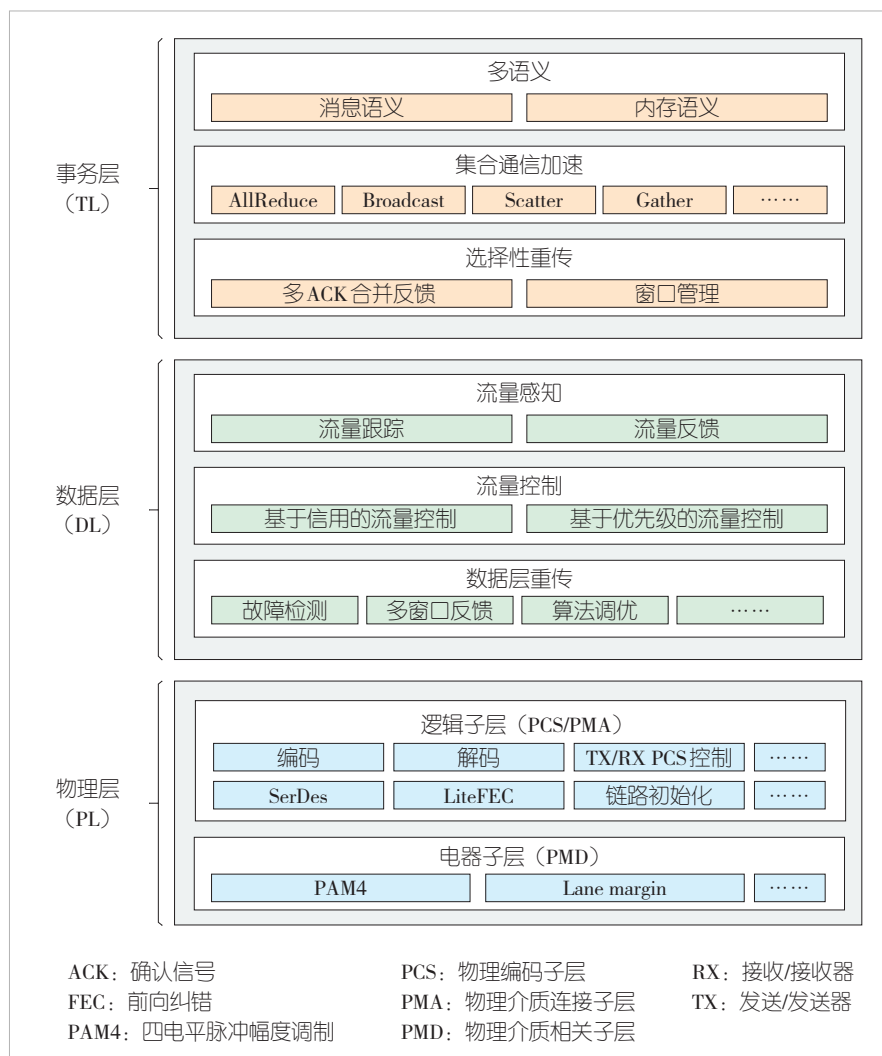


图4 OISA协议栈及各层功能

### 3 同园区机间互联网络

#### 3.1 同园区机间互联网络需求与挑战

智算中心机间网络主要用于承载AI模型训练业务的PP和DP交互的权重、梯度等数据，GPU算力性能高，接入网络带宽100 Gbit/s以上，并通过RDMA来降低传输时延。因此，大模型训练的智算中心网络需要缩短迭代过程中通信传输数据的时间，降低通信开销，从而减少GPU的计算等待，提升计算效率。同时，机间GPU通信流量呈现周期性、同步突发的特点。在大模型训练过程中，通信具有非常强的周期性，且每轮迭代的通信模式均保持一致。在每一轮的迭代过程中，不同节点间的流量保持同步，同时流量以On-Off的模式进行突发式传输。一旦网络出现拥塞或故障，可能会导致数据丢失或传输延迟，从而影响AI模型的训练时效。以上通信流量的特点对机间智算中心网络提出了高带宽、高

吞吐、低延时、高可靠的要求。

同园区智算中心GPU之间通过交换机CLOS组网架构实现互联，每个GPU接入网络带宽需求在100~400 Gbit/s之间，且同步并发要求CLOS组网1:1不收敛。随着GPU性能的增强和GPU集群规模的增加，同园区智算中心GPU对网络的单端口带宽、节点间的可用链路数量及网络总带宽提出了很高的需求。

基于RoCE协议的智算中心网络，通常采用五元组哈希实现链路负载分担技术，并通过PFC、显式拥塞通告(ECN)协议实现网络无损。然而，这种方案在智算中心网络应用中面临4个挑战<sup>[5]</sup>：

挑战1：传统基于逐流的等价多路径路由(ECMP)负载均衡技术，在智算中心网络单流量规模大、流量数量小的情况下会失效，导致流量在交换网络发生极化。这不仅造成链路负载失衡，还会导致部分物理链路因承载过量流量而产生拥塞，最终引发报文丢弃问题，严重影响网络传输效率与稳定性。

挑战2：当前流量进入网络时，在不考虑出端口转发能力的情况下，流量会以“推”的方式进入网络。分布式训练的多对一通信模型会产生大量的Incast流量，造成设备内部队列缓存瞬时突发，进而引发拥塞甚至丢包。PFC和ECN都是

拥塞产生后的被动拥塞控制机制，它们无法从根本上避免拥塞，且1个比特的ECN信号仅能定性表示网络产生拥塞，无法定量表示拥塞程度。对此，端侧需要探测式地调整发送速率，但收敛速度慢，容易导致网络吞吐性能下降。

挑战3：由于网络故障无法避免，因此高效的故障管理和快速的业务恢复机制至关重要。智算中心应配备先进的故障预测工具与自动化故障恢复系统。这些系统可以基于历史数据和实时性能指标，预测并识别潜在的故障点。一旦检测到故障，自动化的故障解除和业务恢复流程就可以迅速启动。这样可以减少系统的停机时间，确保业务的连续性。

挑战4：当前AI超大模型的参数已达千亿甚至万亿级别，训练大模型需要超高算力，并且对显存需求也很高，而更高的显存消耗意味着需要千卡/万卡GPU才能完整存储一个模型的训练过程。组网规模的大幅增长导致网络管理更加

复杂, 拥塞控制、负载均衡的难度也会增加。

### 3.2 同园区机间网络技术业界发展情况

传统智算中心网络技术主要包括 InfiniBand 和 RoCE 两种技术路线。其中, InfiniBand 使用专有的硬件和优化协议, 产业生态链成熟, 但是生态封闭, 网络的建设和维护成本相对较高; RoCE 基于标准以太网, 可以利用现有的以太网基础设施, 降低了网络建设的成本, 但性能受到制约。随着 AI 大模型的快速发展, 智算中心网络技术已经成为全球人工智能巨头关注的焦点, 其核心是新一代以太网技术的突破。

2023 年 5 月, 中国移动携 50 余家业界伙伴发布了全调度以太网技术 (GSE)<sup>[6]</sup>, 原创提出新型以太网转发和调度机制, 以国产技术和产业生态为主, 构造低时延、零丢包的无损网络。当前智算中心服务器主要有两类: 一类是 GPU 集成网卡, 典型产品如华为昇腾 910 系列; 另一类是配备独立网卡的 GPU 服务器, 典型产品如英伟达 H800 系列, 通常需要不同的网络解决方案。GSE 采用统一设计理念, 形成了 GSE-N2N 和 GSE-E2E 两大模式, 以满足各种智算中心的网络需求。其中, GSE-N2N 技术方案适用于 GPU 集成网卡场景, 将网络设备作为全调度网络处理节点 (GSP) 和全调度交换网络 (GSF) 节点, 支持 GSE 的全部功能, 使得 GPU 服务器与 GSE 网络之间无须直接联动, 就能实现天然解耦, 同时确保无损、高性能的集群互联; GSE-E2E 技术方案则适用于配备独立网卡的 GPU 服务器, 通过将部分 GSE 能力延伸至 GPU 服务器的网卡, 将网卡作为 GSP 节点、网络设备作为 GSF 节点, 借助端网协同实现高性能的集群互联。

2023 年 7 月, 博通、思科、Arista、微软、Meta 等美国巨头企业牵头成立超以太网联盟 (UEC)。该联盟秉持生态开放的理念, 将技术路线重心放在以太网通信协议栈的升级上, 通过优化物理层、链路层、传输层和软件应用程序编程接口 (API) 层等方面的技术, 提升以太网在高性能计算和网络通信方面的能力, 其目标是打造面向 AI 时代的超大规模新型网络系统。UEC 的目标愿景和技术理念与中国移动的 GSE 一致。

### 3.3 GSE 核心技术与创新点

GSE 革新以太网底层转发机制和上层协议栈, 创新提出报文容器 (PKTC) 的多路径喷洒、动态全局调度队列 (DGSQ) 的主动拥塞避免等关键技术, 突破无损以太性能瓶颈。主要创新点包括:

#### 1) 基于报文容器的负载均衡机制提升网络吞吐

在智算中心流数少、单流量大、多流并发的场景下, 传统基于流负载分担的方式存在网络拥塞和吞吐率低的问题。GSE 通过在源 GSP 节点将数据包逻辑组装成固定长度的报文容器, 并为每个报文打上 GSE 头、容器标识等。源 GSP 和中间 GSF 以报文容器粒度进行多路径负载均衡, 目的 GSP 节点以报文容器为粒度进行重排, 实现报文保序交付, 使网络吞吐率高达 99%, 如图 5 所示。

#### 2) 基于 DGSQ 的请求授权机制避免拥塞

智算中心存在多打一 Incast 场景, 传统被动拥塞控制有一定的滞后性, 难以有效避免丢包和时延增加。GSE 通过动态调度授权机制避免拥塞, 调度授权流程如图 6 所示, 在源 GSP 接收报文后, 向该报文对应的目的 GSP 进行信令申请; 目的 GSP 收到各个源 GSP 的信令申请后, 按照周期或按次进行信令分配, 在极致负载情况下, 主动控制进入网络的流

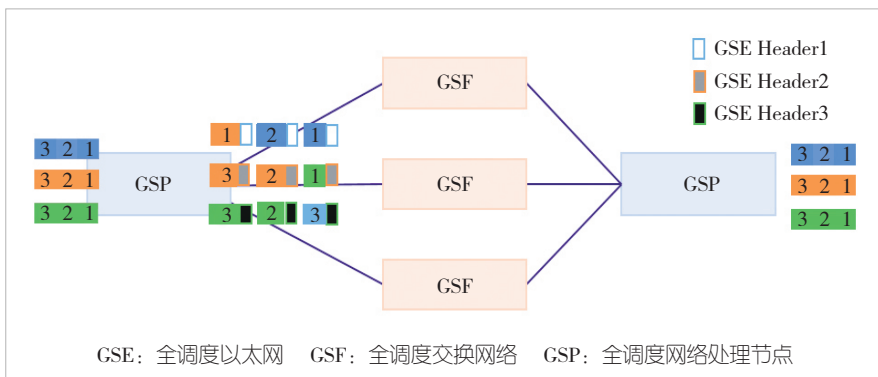


图5 报文容器转发示意图

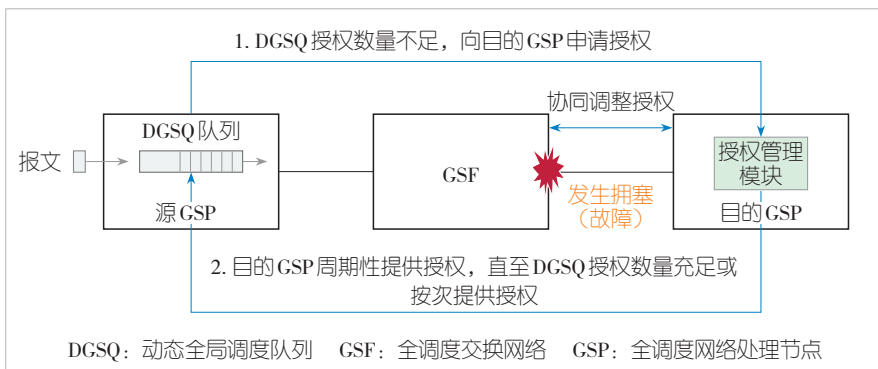


图6 DGSQ调度授权流程

量，实现网络原生无损，并根据实际流量特征按需创建授权队列资源，支持大规模扩展。

### 3) 全局+分布式控制面实现网络设备即插即用

全调度以太网控制面分工如图7所示。集中式全调度操作系统（GSOS）用于全局信息编址（例如设备节点ID等）、设备IP地址配置等。设备端网络侧操作系统（NOS）具备独立的控制面和管理面。GSP/GSF通过eBGP路由打通网络连接，并通过多协议边界网关协议（MP-BGP）在设备间交互控制信息。各设备分布式构建全局转发表，实现GSE报文转发、运行容器负载均衡、DGSQ调度等GSE网络功能，并通过控制器集中式管理和设备分布式控制，实现网络设备的即插即用。

### 4) 微秒级感知网络故障并收敛保证无损

基于报文容器的负载均衡与DGSQ的请求授权机制在正常网络环境下可实现零丢包转发，但当网络发生链路中断或设备故障时，拓扑结构的非对称化仍可能引发局部拥塞。为此，GSF通过光电信号通断检测与误码率实时监测实现微秒级故障感知，并基于66B原子码块或以太报文进行全局通告，将故障通告报文传递至所有关联的GSP。各GSP收到故障报文后进行负载均衡收敛并调整授权，从而保证网络在故障场景下尽力无损，如图8所示。

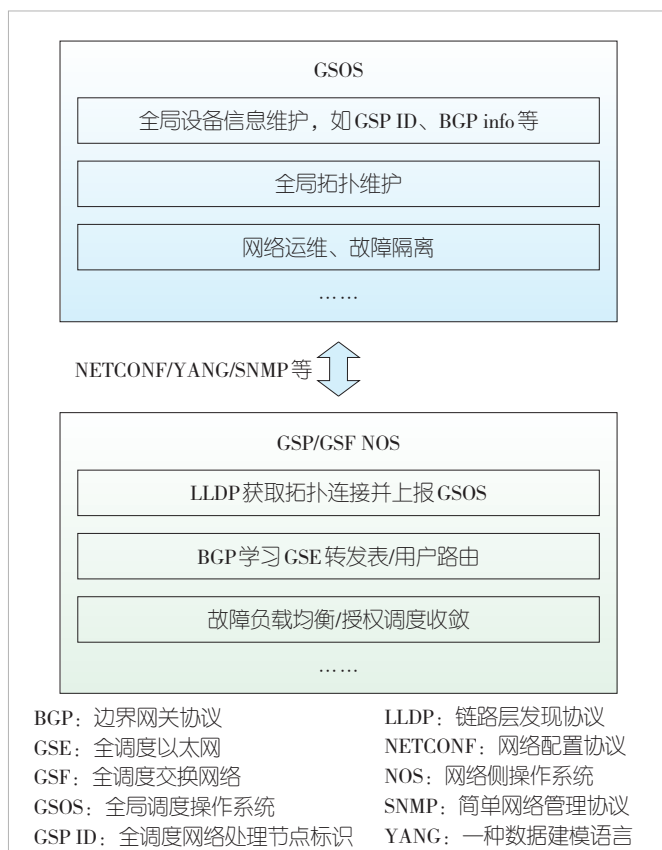


图7 全调度以太网控制面分工

## 3.4 同园区机间网络组网建议

根据园区组网规模和机房条件，园区内单集群建设分为同机房、同楼宇跨机房、同园区跨机房3种场景。为减少对GPU集群算力效率的影响，并简化工程部署设计和运行管理的复杂度，所有GPU服务器都尽量部署在同一个机房内。当基础设施条件不能满足工程部署要求时，可把一个智算集

群分散到多个机房进行跨机房部署，按照机房容量从大到小的原则，充分利用单个机房的容量资源，尽量减少跨机房部署的服务器和网络数量，缩短拉远的布线距离。当在同一个楼宇内找不到满足工程部署条件的机房时，需要考虑采用同园区跨楼宇的拉远部署方式——在园区内的不同楼宇间预埋

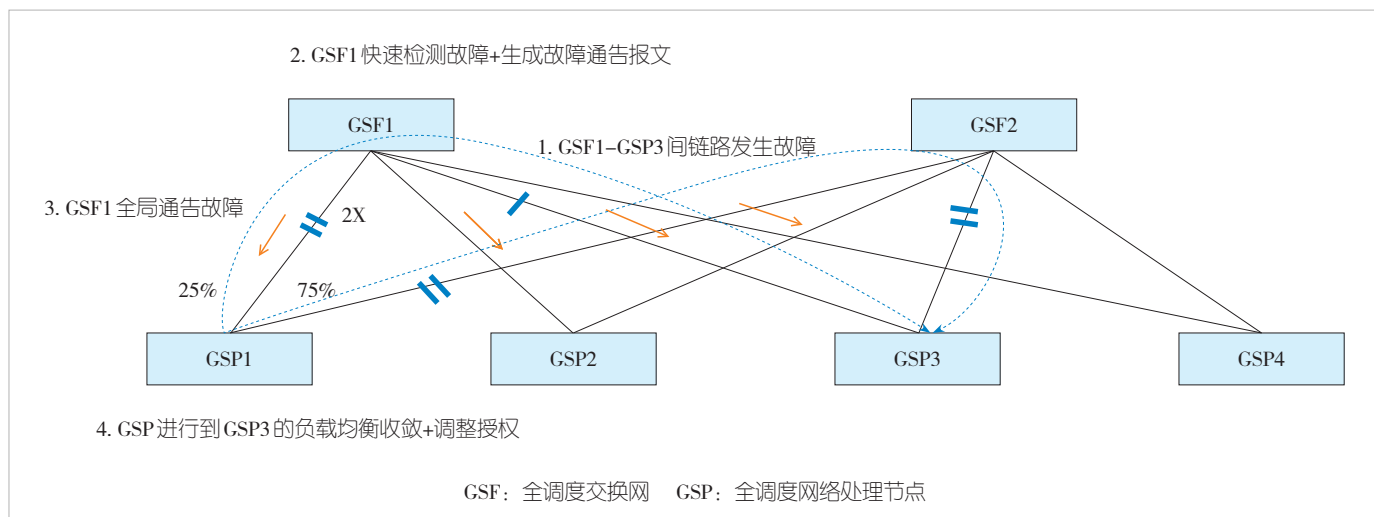


图8 网络可靠性保证流程图



好相应数量的裸纤，以便进行拉远部署。

考虑到交换机转发芯片容量和建设节奏，建议按照POD为单位建设，每个POD接入数千卡至万卡规模，通过平滑扩展POD建设超万卡/十万卡规模的集群。POD间通过高速交换网互联，POD内和POD间根据业务需求可设置一定的收敛比。PP通信量小，但是对时延敏感，在带宽收敛比大、静态时延小的情况下优先跨POD。DP对时延不敏感，但是并行通信量大，在带宽收敛比小、静态时延高的情况下优先跨POD。DP跨POD时可借助训练框架和集合通信算法感知POD，通过分层AllReduce算法进行优化，以减少跨POD间的通信量。

## 4 跨园区智算中心互联网络

### 4.1 跨园区智算中心互联网络需求与挑战

AI集群算力需求的爆发式增长，催生了以网补算的协同训练需求。为突破单中心算力瓶颈，业界开始探索分布式智算中心协同训练模式，将训练任务分散至多个中心并行计算，以扩大整体算力规模。同时，为提升智算中心资源利用效率，服务商期望整合多中心资源，实现AI算力的集中管理与资源池化。然而，跨智算中心互联打破了单体集群的诸多硬件和组网设计前提，导致带宽、时延等网络指标恶化，给网络带来了极致性能挑战。

此外，海量AI训练、推理样本的产生，使得数据网络传输需求呈指数级增长。以乘用车自动驾驶训练为例，每辆路测车每天产生5~10 TB的样本数据，且通常要求在24 h内从分布于不同城市的路测基地传输至智驾训练中心。按照500辆路测车计算，其路测样本数据量每年超过600 PB，约相当于6亿部高清电影的数据量。这对网络带宽、延迟和吞吐量提出了更高要求。因此网络必须具备弹性高带宽能力，以应对大规模数据传输的挑战。

智算中心以大量数据为资源，涉及大量数据在入算、算内和算间网络场景的处理和传输。这些数据是企业十分重要的商业资产，一旦被窃听攻击或泄露，将产生难以估计的经济损失。在跨园区智算中心互联场景下，用于传输智间数据资源的高速互联光纤链路以及相关设施暴露在物理环境中，使得数据窃听、攻击风险倍增，网络面临严峻安全挑战。

### 4.2 跨园区智算中心网络业界技术发展

随着智算中心规模的不断扩大以及算力需求的指数级增长，如何实现高效、稳定、低成本的智算互联，成为了行业发展的关键问题。

思科在跨园区智算中心互联领域提出了路由光网络技术方​​案，通过利用标准化的400G ZR/ZR+可插拔光模块，直接从路由器端口输出相干波长信号，实现了IP层与光层的深度融合。这种架构不仅能简化网络设计，降低运营成本，还可满足智算互联场景对高性能和高可靠性的需求。谷歌积极布局多数据中心分布式训练，其多数据中心间距离最近约为24 km，最远约为80 km。预计到2026年，Google将形成4个园区的跨DC智算互联集群。其Gemini1Ultra系列模型通过光交换机实现多数据中心间的高效、低成本互联<sup>[7-9]</sup>，网络延迟和带宽足以支持常用的同步训练模式，在超级Pod内完成模型并行，在Pod间完成数据并行。诺基亚在智算中心互联领域推出了创新的IP与光融合解决方案，针对智算互联场景中多样化的数据传输距离和性能需求提供了灵活的组网方案：在120 km范围内的400G互联场景中，可通过路由器配置灰光或彩光模块实现高效连接；对于更长距离的传输需求，可采用路由器配置400G多跨光模块方案；而对于多段超长距离的传输，则可通过光网络系统的400G光传输技术实现无缝连接。

腾讯与中国移动正在开展IP+光融合的智能互联技术研究，旨在为智算中心提供高效、灵活的网络互联解决方案<sup>[10]</sup>。在该方案中，通过采用高性能可插拔光模块，并将其直接部署在智算中心互联网络设备上，网络设备能够直接输出高性能彩光信号。这些信号可以直接进入光层系统的合波器，形成密集波分复用（DWDM）合波信号，无须依赖光网络中的OTU单元进行信号的复用和映射，从而显著简化了网络架构。此外，针对IP over DWDM技术，腾讯正在推动光器件的小型化与标准化，即微光学模块的发展，以进一步提升系统的灵活性和可扩展性。传统“路由器+光模块+光传送网+DWDM”的复杂网络架构被简化为“IP+光融合”网络。这种方案不仅降低了处理时延，还减少了多组电层设备的使用频次，使传输路径更加简洁、部署成本显著降低，同时大幅简化了系统部署流程，为智算互联场景提供了高效、低延迟、低成本的基础设施支持。

阿里云在智算中心互联领域通过自研的eCore网络架构，采用单栈单片设计（基于IPv6/SRv6协议栈和单芯片白盒路由器）和多平面架构，实现了全网服务化的技术突破<sup>[11]</sup>。eCore深度融合IPv6/SRv6协议栈与白盒硬件，结合服务化设计，显著提高了网络资源利用率，降低了业务部署复杂度。该架构为智算互联场景提供了高带宽、低延迟的网络基础设施支持，能够高效满足大规模AI训练、分布式计算以及全球化业务的需求，助力智算中心实现高性能、高可靠的网络互联，同时推动智算互联技术的快速发展和规模化应用。

综上,面向智算互联场景,IP+光融合技术日益受到业界关注,并逐渐成为解决高带宽、低延迟网络需求的关键方向。业界在这一领域已开始逐步应用,通过将IP层与光层深度融合,显著提升了网络性能和效率。

### 4.3 跨园区智算中心互联创新关键技术

随着数据流量的爆炸式增长、新兴应用的涌现以及对网络性能要求的不断提升,传统网络架构面临巨大挑战。在此背景下,智算集群互联技术组合应运而生,为构建高效、智能、可靠的网络基础设施提供全新机遇。IP与光通信技术的深度融合,旨在充分发挥两者优势,实现高性能、高带宽、低延迟、高安全和高可靠的网络,以满足日益增长的数字化需求。面向智算互联场景,中国移动在IP广域网络技术上全面创新,革新现有流量分发、拥塞控制和安全加密机制,提出弹性以太网聚合FlexPipe、微流级拥塞控制以及物理层安全加密PHYSec等关键技术,助力构建大运力、无阻碍、高安全的跨智算中心互联网络。主要创新点包括:

#### 1) 弹性以太网聚合FlexPipe技术

当前智算中心互联的路由器设备速率仅为100 Gbit/s,当GPU面临超大AI训练流时,数据中心互联路由器的100G端口必然会出现拥堵丢包的情况。传统广域网采用五元组Hash负载分担方式,常常会将大象流分配至单一链路,即便其他链路仍有空闲资源,该链路也会不可避免地出现拥塞状况,进而产生丢包现象。基于报文组调度的弹性以太网聚合FlexPipe技术能够依据用户的带宽需求,灵活地实现不同带宽速率的以太网端口无损转发。在不依赖特定硬件的情况下,单设备最大可达成3.2T的超宽通道。此外,它还能在超大的FlexPipe管道内部,为多条物理端口间的大象流提供负载分担能力,进一步提升了网络对高带宽突发流量情况的带宽灵活适应能力,为智算中心间的流量无损转发提供坚实保障。

#### 2) 微流级拥塞控制技术

在通过跨智算中心的分布式训练或存算分离等方式进行模型训练时,智算中心资源池间距离的拉远会导致传输时延增加,从而影响网络状态反馈的及时性。然而,智算业务具有超大流、同步传输、任务式、丢包敏感等特征,在进入广域网时,对网络传输的可靠性提出了极高要求。特别是在丢包问题上,智算业务表现出极强的敏感性。因此,需要构建具备稳定连接能力的网络基础设施,以确保智算业务的高效运行。微流级拥塞控制技术能够根据实时网络状况动态调整流控策略,实现流量峰值速率流级别的独立控制和精准反压。相较于传统的PFC机制,微流级拥塞控制技术解决了传

统PFC的头阻、反压风暴和死锁问题,同时基于精准的IP大象流识别和预测,实现了对网络中每一条流的独立监控与动态调整,通过以数据流为单位的精准流量控制并结合大缓存设计及调度,可实现长距网络环境下的业务“零”丢包,将拥塞和故障带来的影响最小化。该技术能够有效解决由短暂拥塞导致的潜在问题,确保数据传输的连续性和完整性。

#### 3) 物理层安全加密PHYSec技术

现有的媒体访问控制安全(MACSec)加密机制是基于数据链路层的安全加密技术,可以保证以太网设备之间数据帧的安全性。但加密后的数据帧仍会暴露以太帧头部信息,同时不能对优先级流量控制帧或Pause帧进行加密。此外,MACSec加密认证过程也会引入较大的封装开销,占用一定的业务带宽。对此,中国移动原创提出的全球首个以太网物理层安全PHYSec技术,通过将传统密码学的理念与以太网物理层技术相融合,提供链路级加解密能力与通道级加解密能力。如图9所示,链路级加解密技术将多个对齐标识(AM)数据段复合形成复帧比特流,作为PHYSec的最小加解密单元,优先在模块内部署。通道级加解密技术针对64B/66B码块流进行全部加密,增加D码块承载解密所需的参数,需要在物理层芯片内部署。该技术具备低时延、高安全、低开销和协议透明等优势,满足数据链路层及所有上层协议的信息防护要求。

## 5 结束语

随着生成式人工智能的蓬勃发展,新型智算网络技术已成为全球创新焦点。智算中心网络是一个多要素融合的复杂系统,是算网的深度融合。中国移动携业界伙伴创新提出OISA、GSE技术体系,以及弹性以太网聚合、精细化拥塞控制、物理层安全等多项创新技术,旨在构建一个标准开放的高性能智算网络技术体系,助力人工智能生成内容(AIGC)业务快速发展。

中国移动已发布《全向智感OISA核心规范》Gen1.1版本协议,GSE全套技术标准《GSE1.0算网协同技术标准》《GSE2.0端网协同技术标准》《GSE2.0网络侧优化技术标准》,以及《新型智算中心以太网物理层安全(PHYSec)架构白皮书》。OISA可支持128张GPU互联,点对点带宽达到800 Gbit/s,极大提升了国产GPU的互联能力。GSE试点表明,OISA的组网性能比传统RoCEv2交换机提升50%以上,达到国际先进水平。

未来,中国移动会持续丰富OISA、GSE协议体系,加速跨园区互联技术创新,并推动构建产业化生态。技术标准的落地依赖AI业务、机内/外交换芯片、网卡、网络设备等

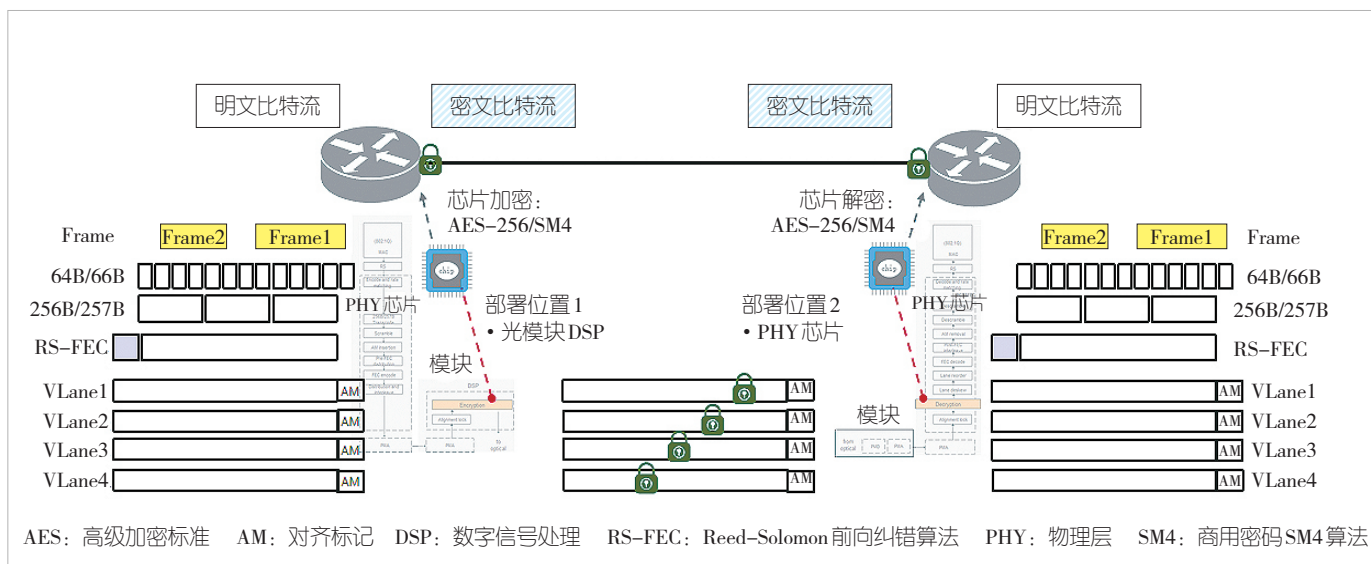


图9 PHYSec技术架构

上下游的协同创新。由于创新难度大、开发周期长，我们希望产业各界能够携手合作，协同创新，共同推动技术进步，为产业发展贡献力量。

[11] 阿里云基础设施. 阿里云基础设施网络 2024 年创新总结 [EB/OL]. (2025-01-20) [2025-03-10]. <https://mp.weixin.qq.com/s/JnziKrk8dYcOcafj29apA>

#### 参考文献

- [1] KAPLAN J, MCCANDLISH S, HENIGHAN T, et al. Scaling laws for neural language models [EB/OL]. [2025-03-10]. <https://arxiv.org/abs/2001.08361v1>
- [2] WEI J, TAY Y, BOMMASANI R, et al. Emergent abilities of large language models [J]. Transactions on machine learning research, 2022(8): 1-30
- [3] DU Z X, ZENG A H, DONG Y X, et al. Understanding emergent abilities of language models from the loss perspective [EB/OL]. [2025-03-10]. <https://arxiv.org/abs/2403.15796v3>
- [4] HOWARTH J. Number of parameters in GPT-4 [R]. 2024
- [5] 中国移动通信研究院. 面向 AI 大模型的智算中心网络演进白皮书 [R]. 2023
- [6] 段晓东, 程伟强, 王瑞雪, 等. 面向新型智能计算中心的全调度以太网技术 [J]. 中兴通讯技术, 2023, 29(4): 57-63. DOI: 10.12142/ZTETJ.202304011
- [7] Gemini Team. Gemini: a family of highly capable multimodal models. Google [EB/OL]. [2025-03-10]. [https://storage.googleapis.com/deepmind-media/gemini/Gemini\\_1.0.pdf](https://storage.googleapis.com/deepmind-media/gemini/Gemini_1.0.pdf)
- [8] HONG C Y, MANDAL S, AL-FARES M, et al. B4 and after: managing hierarchy, partitioning, and asymmetry for availability and scale in Google's software-defined WAN [C]//Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication. ACM, 2018: 1-44. DOI: 10.1145/3230543.3230545
- [9] POUTIEVSKI L, MASHAYEKHI O, ONG J, et al. Jupiter evolving: transforming Google's datacenter network via optical circuit switches and software-defined networking [C]//Proceedings of the ACM SIGCOMM 2022 Conference. ACM, 2022: 66-85. DOI: 10.1145/3544216.3544265
- [10] CAICT 算力. MegaScaleOut: 聚焦百万量级 GPU 集群, ODC 启动算力网络新项目 [EB/OL]. (2024-12-17) [2025-03-10]. [https://mp.weixin.qq.com/s/k\\_7xtw-1uMUbK7\\_Km1hcrw](https://mp.weixin.qq.com/s/k_7xtw-1uMUbK7_Km1hcrw)

#### 作者简介



**段晓东**, 中国移动通信有限公司研究院副院长, 教授级高工, 享受国务院特殊津贴专家, 工信部通信科技委信息通信专家组专家, “新世纪百万人才工程”国家级人选; 长期从事 AI 智算网络、下一代互联网、5G 网络架构、SDN/NFV 下一代网络及 6G 网络研究等工作, 先后主持了多项国家重大专项课题, 并获得国家科技进步奖特等奖等多个奖项; 获授权专利 50 余项。



**程伟强**, 中国移动通信有限公司研究院基础网络技术研究所副所长, 教授级高工; 长期从事下一代互联网、数据中心网络、传输网等方面的技术研究和标准推动工作。



**张昊**, 中国移动首席专家、中国移动通信有限公司研究院网络与 IT 技术研究所所长, 高级工程师; 长期从事 4G/5G 核心网、边缘计算、算力网络、新型智算等 ICT 融合领域技术研发工作。



# 智算中心组网技术及应用



## Intelligent Computing Center Networking Technology and Applications

段威/DUAN Wei<sup>1,2</sup>, 于浩/YU Hao<sup>2</sup>, 李和松/LI Hesong<sup>1,2</sup>,  
张征/ZHANG Zheng<sup>1,2</sup>, 王宁诚/WANG Ningcheng<sup>1,2</sup>

(1. 移动网络和移动多媒体技术国家重点实验室, 中国 深圳 518055;

2. 中兴通讯股份有限公司, 中国 深圳 518057)

(1. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China;

2. ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.202502009

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250508.1356.002.html>

网络出版日期: 2025-05-08

收稿日期: 2025-03-05

**摘要:** 智算中心网络架构正面临高性能、低延迟、高扩展性和可靠性的挑战。当前多种主流的智算中心组网拓扑结构各具优缺点。针对智算中心各种组网方案, 边界网关协议 (BGP)、开放式最短路径优先 (OSPF)、中间系统到中间系统 (IS-IS)、胖树路由协议 (RIFT) 等具有不同适用性。中兴通讯在智能计算中心网络架构与部署方案领域取得突破性进展, 创新性地提出了 Scale-Up 与 Scale-Out 融合组网解决方案, 并通过协议优化提升等价多路径路由 (ECMP) 路径选择精度与故障收敛效率。这些研究和实践可为智算中心网络设计提供参考, 推动智算中心技术的发展和应用。

**关键词:** 网络协议; 组网技术; 智算中心网络协议; 融合组网

**Abstract:** The network architecture of intelligent computing centers is facing challenges in terms of high performance, low latency, scalability, and reliability. Currently, there are several mainstream networking topologies for intelligent computing centers, each with its own advantages and disadvantages. Meanwhile, for various networking solutions to intelligent computing centers, protocols such as border gateway protocol (BGP), open shortest path first (OSPF), intermediate system to intermediate system (IS-IS), and routing in fat trees (RIFT) have different applicabilities. ZTE Corporation has made breakthrough progress in intelligent computing center network architecture and deployment solutions, innovatively proposing a converged networking solution integrating both Scale-Up and Scale-Out approaches. By employing protocol optimizations, the precision of equal-cost multi-path (ECMP) path selection and the efficiency of fault convergence have been significantly enhanced. These studies and practices can provide references for the network design of intelligent computing centers and promote the development and application of intelligent computing center technologies.

**Keywords:** network protocol; networking technology; intelligent computing center network protocol; converged networking

**引用格式:** 段威, 于浩, 李和松, 等. 智算中心组网技术及应用 [J]. 中兴通讯技术, 2025, 31(2): 63-71. DOI: 10.12142/ZTETJ.202502009

**Citation:** DUAN W, YU H, LI H S, et al. Intelligent computing center networking technology and applications [J]. ZTE technology journal, 2025, 31(2): 63-71. DOI: 10.12142/ZTETJ.202502009

近年来, 大模型技术迅猛发展<sup>[1-5]</sup>, 模型能力快速增强。大模型技术已经深入到人们生产生活的方方面面, 人类社会正在快速进入人工智能 (AI) 时代。参数规模是影响模型能力的关键要素, 已经从千亿级别向十万亿级别快速增加。

由于未来需要支撑 10 万亿级别参数规模, 图形处理器 (GPU) 的需求急剧增加。根据摩根士丹利的研究报告, 生成式预训练变换器 GPT-5 至少需要 25 000 个 GPU。为了支撑如此大规模 GPU 卡的应用, 智算中心组网技术已经成为业界研究的热点<sup>[6-7]</sup>。

## 1 智算中心网络拓扑和协议

### 1.1 智算中心主流网络拓扑介绍

智算中心的网络设计需满足高性能、低延迟、高扩展性

和可靠性的需求。不同业务对网络性能 (带宽、延迟、吞吐量等) 的要求不同, 因此需要选择合适的网络设计。例如, 智算网络需要低延迟和高带宽, 而云服务则需要高扩展性和弹性。此外, 数据中心的规模从几十台服务器到数十万台不等, 网络设计需满足不同的扩展需求。目前, 主流的智算中心组网拓扑结构包括 CLOS、Dragonfly、Torus。每种网络拓扑有各自的特点及应用场景, 在标准组网下也衍生了一些变种组网。

#### 1.1.1 CLOS 网络

CLOS 网络采用多级分布式架构, 分为接入层 (Leaf 层)、汇聚层 (Spine 层) 和核心层 (可选), 具有分层结构、无阻塞性、模块化设计和高带宽利用率等特点。CLOS 网络通过多路径设计实现流量无阻塞转发, 适用于大规模通信负

载。其模块化设计易于扩展，只需增加交换机和链路即可扩展网络规模。

### 1) 两层CLOS架构

如图1所示，假设交换机端口为 $n$ ，则两层CLOS最大规格可支持 $n^2/2$ 个GPU卡互联，所需交换机的数量为 $N+N/2$ 。

以128口400 Gbit/s的交换机、连接带有8个GPU和8个网口的服务器为例，组网按照无收敛比来设计。单个交换机上行64端口连接Spine交换机，下行64端口可连接8台服务器（每台包含8块GPU卡），单个交换机与8台服务器组成一个机架顶交换机（ToR）区域。Spine层有64个交换机，共连接128个ToR区域。其中，Spine1交换机的128个端口连接每个ToR交换机的1号口，Spine2交换机的128个端口连接每个ToR交换机的2号口。依次类推，Spine64交换机的128个端口连接每个ToR交换机的64号口。两层组网组成一个Pod，该Pod拥有8 192个GPU和192台交换机。

### 2) 三层CLOS架构

三层CLOS架构是在二层架构的基础上增加Core层交换机而形成的。通过Core层交换机，三层CLOS架构能够实现Spine层交换机的互联。假设交换机端口为 $n$ ，则三层CLOS最多可支持 $2 \times (n/2)^3$ 个GPU卡互联，所需交换机的数量为 $N^2+(N/2)^2$ 。

以128口400 Gbit/s的交换机、连接带有8个GPU和8个网口的服务器为例，组网按照无收敛比设计。对于ToR区域，与两层组网相同，单个Leaf交换机下挂8机共64卡GPU。对于Spine层，每个交换机需要有一半端口向上连接第三层，因此剩余64下联口向下连接64个ToR，以便组成一个集群Pod。其中，每个Pod拥有4 096个GPU卡、128台

交换机、4 096个上联口。

第三层Core交换机共有128个下联口连接Pod，每个交换机连接每个Pod中具有相同编号的上联口。例如：Core1交换机连接Pod1~Pod128中Spine1的1口，Core2交换机连接Pod1~Pod128中Spine1的2口，依此类推，Core4096交换机连接Pod1~Pod128中Spine64的64口。因此，Core层共需要4 096台交换机。三层组网可支持 $2 \times (128/2)^3=524\ 288$ 个GPU，共需要20 480台交换机。

### 1.1.2 多轨CLOS网络

多轨CLOS网络在传统CLOS的基础上增加了多轨设计，支持更高的并发通信性能和带宽利用率。其特点包括多轨设计、低延迟、高容错和简化布线，适合超大规模网络。多轨设计通过独立的数据路径实现流量的分布式转发，显著提高了带宽利用率和容错能力。典型的组网应用如阿里的高性能网络（HPN）<sup>[8]</sup>、字节的超大规模网络Megescale<sup>[9]</sup>、百度的自研数据中心等<sup>[10]</sup>。

#### 1) 两层多轨CLOS架构

如图2所示，假设交换机端口为 $N$ ，单台服务器GPU数为 $M$ （ $M < N$ ），则两层CLOS多轨可支持 $2 \times (N/2)^2 \times M$ 个GPU卡互联，所需交换机的数量为 $(N+N/2) \times M$ 。

以128口400 Gbit/s的交换机、连接带有8个GPU和8个网口服务器为例，组网按照无收敛比设计。在多轨CLOS架构的设计中，由于每个服务器的同号GPU与同一个交换机连接，因此对于一个ToR来说，8卡服务器就需要8台交换机（Leaf1~Leaf8），每台交换机64口上联Spine层，64口下联服务器网卡。其中，Leaf1交换机1口接服务器1的GPU1，

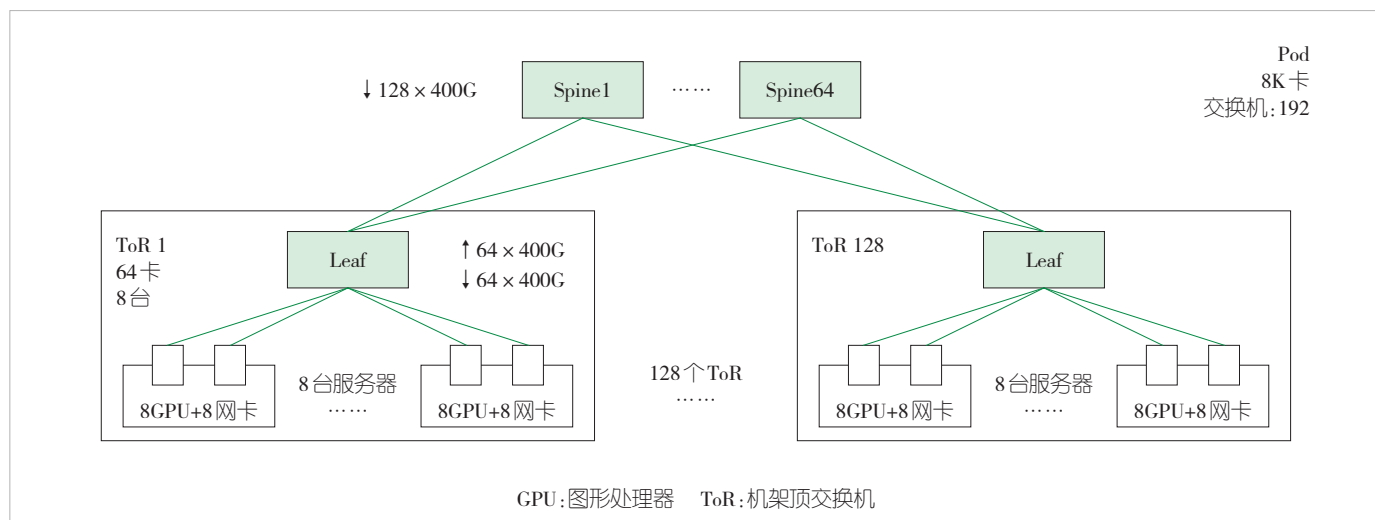


图1 两层CLOS组网架构

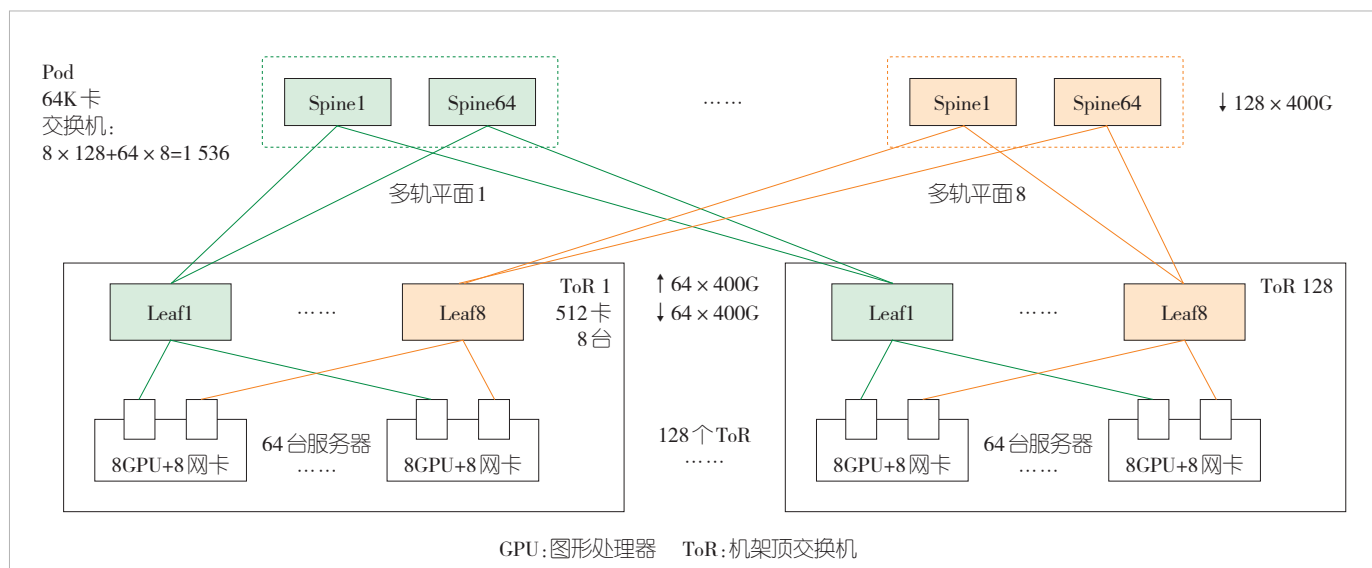


图2 两层多轨CLOS组网架构

2口接服务器2的GPU1，依此类推，64口接服务器64的GPU1；Leaf2交换机连接所有服务器的GPU2，Leaf8连接所有服务器的GPU8。因此，单个ToR共有8台交换机，64台服务器共有512卡。

网络第二层也是按平面进行划分的，共有8个平面。对于单个平面而言，其与传统CLOS网络相同，Leaf层交换机上联64口分别接Spine层64台交换机，Spine层每个交换机128口可以连接128个ToR。两层网络组成一个Pod，共有 $2 \times (128/2)^2 \times 8 = 65\,536$ 个GPU卡，交换机数量为 $(128+128/2) \times 8 = 1\,536$ 台。

## 2) 三层多轨CLOS架构

三层多轨CLOS架构增加Core层交换机，使Spine交换机按照轨道划分进行互联。假设交换机端口为 $N$ ，单台服务器GPU数为 $M$  ( $M < N$ )，则三层多轨CLOS可支持 $2 \times (N/2)^3 \times M$ 个GPU卡互联，所需交换机数量为 $[N^2 + (N/2)^2] \times M$ 。

以128口400 Gbit/s的交换机、连接8GPU 8网口服务器为例，组网按照无收敛比设计。三层CLOS多轨ToR与两层相同。Spine层交换机需要有一半端口上接Core层，下接的ToR区域数量为两层网络的一半，即64个ToR区域。因此，单个Pod可连接32 768个GPU卡、1 024台交换机。

第三层网络与第二层网络类似，也采用分平面方式进行组网，共有8个平面，且每个平面的第三层交换机共有4 096台。该层网络的连接方式与普通CLOS网络一样，相同编号的GPU连接在同一个交换机上。因此，三层多轨CLOS网络最大可支持 $2 \times (128/2)^3 \times 8 = 4\,194\,304$ 个GPU卡，需要的交换机数量为 $(128^2 + (128/2)^2) \times 8 = 163\,840$ 。

## 1.1.3 Dragonfly 网络

Dragonfly是当前应用最广泛的直连拓扑网络架构<sup>[11]</sup>。本地通信通过组内交换机完成，组间通信通过高效的全局链路完成，减少了传统分层结构中的多跳路径。Dragonfly的拓扑结构分为3层：Switch层、Group层、System层。

Switch层：包含1个交换机及其相连的 $p$ 个计算节点。

Group层：包含 $a$ 个Switch层，这 $a$ 个Switch层的 $a$ 个交换机是全连接(All-to-all)的。换言之，每个交换机都有 $a-1$ 条链路分别连接到其他的 $a-1$ 台交换机。

System层：包含 $g$ 个Group层，这 $g$ 个Group层也具有全连接的特点。

对于单个Switch交换机，它有 $p$ 个端口连接到计算节点， $a-1$ 个端口连接到Group内其他交换机， $h$ 个端口连接到其他Group的交换机。因此，我们可以通过计算得到网络的如下属性：

- 每个交换机的端口数为 $k = p + (a-1) + h$ ；
- Group的数量为 $g = ah + 1$ ；
- 网络中计算节点数共有 $N = ap(ah+1)$ 个；
- 如果我们把一个Group内的所有交换机视为一个总交换机，那么这个交换机的端口数为 $k' = a(p+h)$ 。

## 1.1.4 Dragonfly+网络

Dragonfly+网络是在Dragonfly网络基础上进一步优化的智算中心网络和数据中心网络架构<sup>[12]</sup>，它将Group内的通信结构变为了Fat-tree。相较于传统的Dragonfly网络，Dragonfly+主要在拓扑结构、全局链路分布、路由策略以及可扩展性方面进行了改进。Dragonfly+的组网方式与Dragonfly基



本相同，其基本组网结构如下：

节点分组（Pod）：网络被划分为多个组，每组由若干交换机和计算节点组成，组内可采用CLOS架构。

组内通信：组内交换机和节点通过高带宽连接实现低延迟通信。

组间通信：使用分层全局链路连接各组，链路被分为近邻和远程两种。

Dragonfly+组网架构可以灵活分配组内通信和组间通信的连接数量，以适应不同规模的数据中心网络。在最大拓扑结构下，每对组都由一个全局链接连接，从而形成保留Dragonfly属性的最大拓扑。在中型拓扑结构下，每个骨干路由器都通过一个全局链路连接到其他所有组。在小型拓扑结构下，每个脊椎路由器都通过几个并行链路连接到其他组。Dragonfly+最大拓扑组网如图3所示。

组内网络采用Leaf-Spine两层CLOS架构。其中Leaf与Spine采用全连接方案，每个Leaf交换机都连接到组中的 $p$ 个主机和 $s$ 个Spine交换机，每个Spine交换机都连接到同组的 $l$ 个叶子路由器和其他分组的Spine交换机。为了保证组内网络性能，一般建议： $p=l=s=h$ 。假设交换机的端口数为 $k$ ，此时 $p=h=k/2$ 。单个组内的主机数量为 $N=p \times l=k^2/2$ 。

### 1.1.5 Torus 网络

Torus 网络采用网格化设计，通常用于计算芯片直连。典型的Torus网络有Google张量处理器TPU V4<sup>[13]</sup>。Torus网络的节点排列表现为一维、二维或三维结构<sup>[14]</sup>，具体特点如下：

一维Torus：所有节点连接成一个环状拓扑；

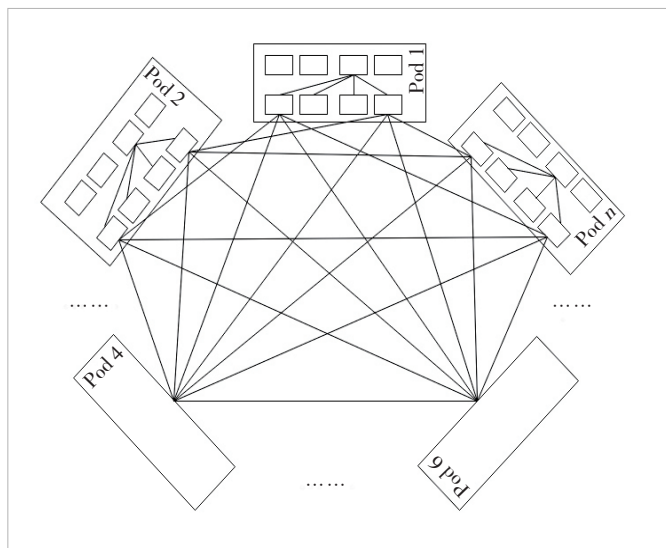


图3 Dragonfly+组网架构

二维Torus：节点排列成二维网格，行列两端节点通过额外链路闭合；

三维Torus：节点扩展为三维网格，每个维度的边界闭合，形成环状结构，如图4所示。

### 1.1.6 网络拓扑对比分析

各类组网架构根据性能需求、延迟要求、扩展性和成本进行选择，以应对不同的数据中心和计算任务需求。表1展示了各种组网架构的优缺点。

## 1.2 智算中心网络协议设计

如何在多种智算中心组网架构上实现高效路由是智算中心设计面临的关键问题之一。传统路由协议模式分为集中式、分布式和混合式。其中，集中式路由是指引入软件定义网络（SDN）控制器，由控制器下发流表并指导交换机转发的方式。分布式路由是指利用动态路由协议，让所有交换机自主交互路由信息。各个交换机自主计算路由，得出整网的无环路由表以供转发使用。这里所指的动态路由协议包括边界网关协议（BGP）、开放式最短路径优先（OSPF）、中间系统到中间系统（IS-IS）、胖树路由协议（RIFT）、路由信息协议RIP等。混合式路由是指，将SDN控制器与分布式路由协议相结合，利用各自的优势进行部署的模式。

在现有智算中心中，主流趋势是采用分布式路由模式。在一些需要精细化控制的场景中，混合式的路由模式也有一定应用。然而，无论是分布式模式还是混合式模式，动态路由协议都是其中的核心组成部分。路由协议的对比如表2所示。

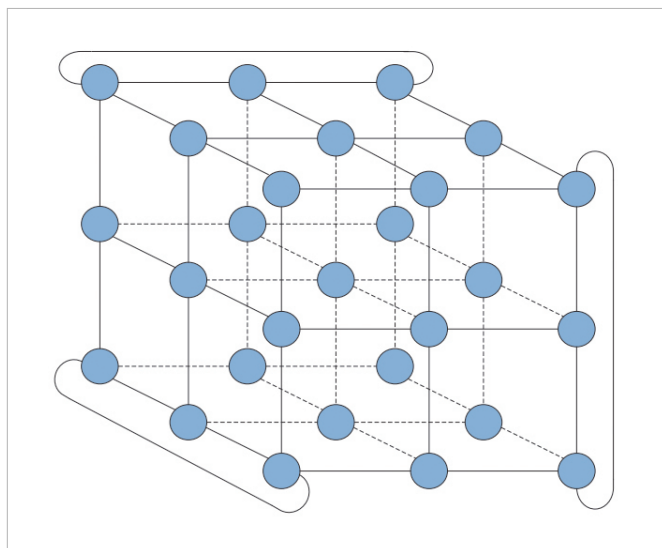


图4 三维Torus组网架构

表1 智算中心网络拓扑对比

网络架构	优点	缺点
CLOS	扩展性强,带宽利用率高,无阻塞性,商用成熟度高	布线复杂,成本和能耗高
Dragonfly	低延迟、高性能,可实现全局通信优化	部署复杂,不适合小规模网络
Dragonfly+	链路利用率更高,灵活性强,超大规模扩展,容错能力强	拓扑复杂,部署成本高,非典型流量模式适应性弱
Torus	低延迟、低成本,拥有路径冗余保护	扩展性差,灵活性低,性能不足
单机多轨	性能更高,容错能力强,布线简化,扩展灵活	布线复杂,成本高,功耗较高

表2 智算中心网络协议对比

内容	BGP	OSPF	IS-IS	RIFT
协议类型	路径矢量协议	链路状态协议	链路状态协议	链路状态+距离矢量
最初设计目标	跨自治系统(AS)互联	企业级内部网络路由	企业级内部网络路由/运营商骨干网	大规模 CLOS/Spine-Leaf 架构
分层能力	基于AS分层	基于Area分层	基于Level分层	基于Level分层
收敛速度	较慢(依赖定时器)	快(泛洪+SPF计算)	快(泛洪+SPF计算)	极快(优化泛洪与计算机制)
扩展性	极高(支持超大规模网络)	中等(Area规划复杂度高)	高(基于Level规划)	极高(专为CLOS架构优化)
策略控制	极强	较弱	较弱	强
支持零接触部署(ZTP)	否	否	否	是
适用架构	大规模平面网络	传统三层架构	CLOS/Fat-tree/运营商架构	CLOS/Fat-tree
典型应用场景	超大规模数据中心	中小型数据中心	运营商/电信级数据中心	新一代数据中心

BGP:边界网关协议    IS-IS:中间系统到中间系统    OSPF:开放式最短路径优先    RIFT:胖树路由协议    SPF:最短路径优先

在数据中心网络设计中，动态路由协议的选择需综合考虑网络规模、拓扑结构、扩展性、收敛速度和运维复杂度等因素。

- 从架构角度考虑：在数据中心选择使用 CLOS/Fat-tree 架构时，可优先考虑 RIFT、IS-IS 或者 BGP 协议。如果是传统的三层架构，仍然可以考虑使用 OSPF 和 IS-IS 协议。
- 从规模角度考虑：如果是小型数据中心（节点数量小于 500），可考虑使用 OSPF 协议；如果是中型数据中心（节点数量为 500~5 000），则可考虑 IS-IS、RIFT 或者 BGP 协议；如果是大型数据中心（节点数量在 5 000 以上），则可考虑 BGP 或者 RIFT 协议。
- 从大规模运维难度考虑：如果管理能力强，可考虑使用 BGP 协议；如果非常需要自动运维的能力，则可以考虑 RIFT 协议。

尽管上述拓扑结构和协议为智算中心组网提供了多样化选择，但在超大规模部署场景中仍面临挑战：CLOS 架构的布线成本与能耗随规模增长呈指数级上升，Dragonfly+ 的复杂拓扑增加了运维难度，传统协议在等价多路径路由（ECMP）路径动态优化方面存在响应滞后问题。此外，当前组网方案普遍采用机内（Scale-Up）与机间（Scale-Out）网络独立部署模式，导致资源利用率低下、运维复杂度激增。针对这些挑战，我们提出面向超大规模智算中心的融合

组网架构与协议创新方案，通过技术整合突破传统架构的局限性。

2 中兴通讯智算中心组网方案和协议创新

当前智算中心商用规模部署的组网方案以 CLOS 为主。该组网方案凭借高性能、低延迟、强扩展性和高可靠性优势，已成为现代智算中心大规模网络设计的重要选择。CLOS 也是当前中兴通讯推荐的主流组网方案。基于该组网方案，我们在网络架构优化、传输效率提升等方面不断进行创新和改进，以更好地满足智算中心日益增长的业务需求。

2.1 Scale-Up 和 Scale-Out 融合组网架构

在现有智算中心网络架构中，Scale-Up 网络和 Scale-Out 网络是两类最重要的后端网络。Scale-Up 网络（如 NV - Link）与 Scale-Out 网络（如 CLOS）通常采用独立部署方案，存在成本、性能和运维三方面核心矛盾。针对这一问题，我们提出 Scale-Up 与 Scale-Out 融合组网架构，通过协议层与拓扑层的协同设计实现资源的统一管理。

Scale-Up 网络的功能定位是构建更大的逻辑 GPU，旨在解决单个计算节点或者超节点内 GPU 之间的高速互联。典型的实现方式是英伟达通过 NVLink/NVSwitch 实现最大 576 个 GPU 之间的 TB 级无阻塞互联。与 Scale-Out 网络相比，Scale-Up 网络具有如下 3 个特征：

### 1) 超高带宽

Scale-Up 网络对带宽的要求比 Scale-Out 高出一个数量级。以 GB200 为例, Scale-Up 网络内 GPU 点到点通信带宽为双向 1.8 TB/s, 对应 Scale-Out 网络带宽为单向 800 Gbit/s。Scale-Up 网络带宽是 Scale-Out 网络带宽的 9 倍。

### 2) 超低延迟

由于需要支持 Load/Store 同步内存语义, Scale-Up 网络是一种延迟敏感型网络, 要求纳秒级的网络延迟。相对而言, Scale-Out 的网络延迟需求通常在微秒级。以 Ualink 为例, 它要求用于 Scale-Up 场景的交换芯片延迟控制在 300 ns 以内。这一指标与当前 NVSwitch 基本持平。

### 3) 有限规模

Scale-Up 网络另外一个重要的特征是组网规模有限。从 AI 基础设施的发展趋势来看, 未来几年 Scale-Up 网络的规模将维持在千卡以下。这与 Scale-Out 网络动辄要求万卡甚至百万卡形成鲜明的对比。

总体而言, Scale-Up 网络侧重性能, Scale-Out 网络侧重规模, 两者密切协同完成一个高效 AI 集群的构建。在典型的 AI 大模型训练和推理过程中, 通常会设计复杂的混合并行策略, 根据并行流量的特征差异综合利用 Scale-Up 和 Scale-Out 的网络资源, 例如: 将对带宽要求极高的张量并行 (TP) 流量和对延迟要求较高的控制流量规划到 Scale-Up 通信域, 将对带宽和延迟不太敏感的流水线并行 (PP) 和数据并行 (DP) 流量疏导到 Scale-Out 网络中。随着 AI 模型参数规模和模型架构的不断演进, 并行策略变得日益复杂, 但这种模型切分策略和网络的最优匹配逻辑却一直存在。

正是由于存在上述明显的差异性, 目前 Scale-Out 和 Scale-Up 被建设为两张独立的网络。这两张网络存在多种技术路线, 异构特征明显。具体而言, Scale-Out 网络的技术路线分为无限带宽 (IB) 和以太网两个封闭的产业生态。其中, IB 以英伟达为主, 而以太网被大部分厂商所支持。Scale-Up 网络的技术实现则更加多样化, 总体而言可以分为总线网络化和网络总线化两大类。其中, 总线网络化是指, 以传统总线型协议 (如 PCIe) 为基础, 通过吸收部分以太网的底层技术来满足 Scale-Up 网络的需求, 典型的实现方式如 NVLink 和 Ualink; 网络总线化是指, 以传统网络协议 (如以太网) 为基础, 通过引入总线协议的设计来满足 Scale-Up 网络的需求, 典型的实现方式如美国的基于以太网的特斯拉传输协议 TTPoE、中国的高通量以太网等。总体而言, 当前 Scale-Up 和 Scale-Out 独立组网仍存在如下缺陷:

#### 1) 建设成本高昂: 两张独立的网络导致资源利用率不

高, 为了达到最优的组合性能, 通常只能从一家供应商购买, 进而导致整体建设成本高昂。

#### 2) 管理运维困难: 运维人员需要同时熟悉两套技术栈, 故障定位、故障恢复难度将成倍增加, 导致运营成本居高不下。

为了解决上述组网痛点问题, 基于对未来智算互联行业趋势的准确把握, 中兴通讯创新性地提出了 Scale-Out 与 Scale-Up 融合组网的理念和对应的产品解决方案, 总体特征如下:

#### 1) 统一互联协议

针对 Scale-Out 和 Scale-Up 场景的需求, 中兴通讯基于以太网技术为基础, 构建了开放互联协议 OLink 技术栈。该技术栈在底层兼容以太网的基础上, 通过物理层和事务层的创新设计, 配合极简的端侧互联网协议 (IP) 设计, 使系统既可以满足 Scale-Out 网络超大规模扩展的需求, 也可满足 Scale-Up 网络高性能互联的需求。

#### 2) 融合极简组网

由于底层协议都被统一到以太网, 在不影响业务性能的前提下, Scale-Up 和 Scale-Out 网络可以融合形成一张物理网络。这样在极致组网场景中可以不再需要独立的数据处理单元 (DPU) 网卡, 从而大幅降低整体组网的复杂性和成本。该组网方式通过灵活的服务质量 (QoS) 策略实现不同业务流量之间的性能隔离。

#### 3) 降低组网成本。当 Scale-Up 和 Scale-Out 融合后, 原来分布在 GPU 和 DPU 的端侧功能被集中到 GPU 侧。此时硬件构成变得更为简单 (不再需要独立的 DPU 网卡和配套的 PCIe Switch), 管理运维也进一步统一, 整体组网成本大幅下降。

图 5 展示了中兴通讯基于 OLink 技术推出的融合组网架构。GPU 集成 OLink 提供的端侧 IP, 与支持 OLink 协议的交换芯片结合, 实现统一的 Scale-Out 和 Scale-Up 通信。在超节点 (典型如 64 卡) 内的 GPU 通过第一级交换节点实现高带宽通信域, 并通过 OLink 协议实现低延迟、高带宽、内存语义以及在网计算等与 Scale-Up 相关的功能。超节点之间通过两级甚至三级交换实现大规模组网, 满足 Scale-Out 通信的基本要求。Scale-Up 和 Scale-Out 网络在 GPU 和第一级交换机上物理共享, 但在协议上存在差异。

基于以太网技术的 Scale-Up 和 Scale-Out 融合组网不是一蹴而就的, 它需要在传统以太网的基础上针对场景需求做大量的技术创新。业界一直都有总线网络化和网络总线化两种探索路径。Ualink 是典型的总线网络化技术路线, 而 OLink 则选择了网络总线化这种具有更强演进能力的技术路线, 两者的对比如表 3 所示。从场景需求来看, Scale-Up 是一个介于片上网络 (NOC) 总线和网络 Fabric 之间的互联场



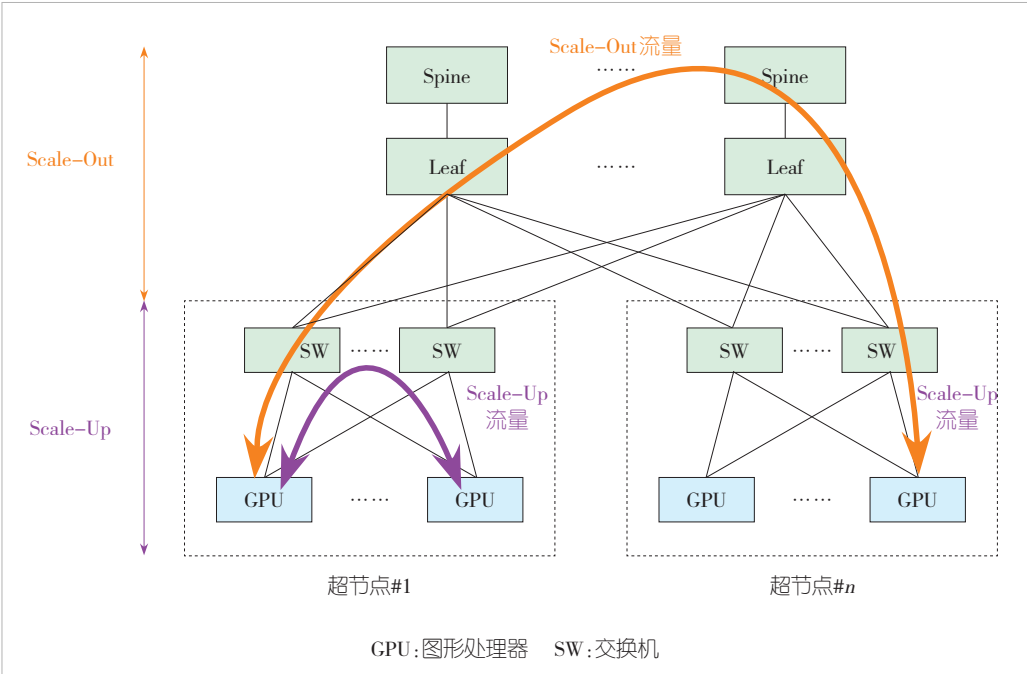


图5 中兴OLink融合组网示意图

景，要求满足TB级超高互联带宽、K级互联规模、原生内存语义、低功耗、纳秒级延迟以及统一内存编址的需求。以太网在互联带宽、互联规模等方面相比于传统总线技术有明显优势，但在其他几个指标维度则存在不足。OLink借助Transaction Packing、芯片PPA（指性能、功耗、面积）优化、低时延前向纠错（FEC）和定制化报文格式等技术措施，充分满足Scale-Up在内存语义、功耗、低延迟以及内存统一编址方面的需求。在技术指标上，OLink与Ualink殊

途同归，但依托以太网快速发展的产业生态，OLink在降低成本、提升兼容性等方面更具优势。

总体而言，OLink是中兴通讯长期基于以太网底层技术创新提出的面向未来AI基础设施的互联解决方案。作为Scale-Up和Scale-Out融合组网的技术倡导者和产品践行者，中兴通讯将依托OLink方案，持续深化技术迭代与生态共建，加速AI基础设施网络架构的升级演进。

2.2 基于BGP扩展的ECMP动态优化

智算网络（如AI训练/推理场景）对网络的低时延、高带宽和零丢包特性提出了极高要求。尽管第1章所述的RIFT协议在CLOS网络中实现了快速收敛，但其ECMP路径选择仍然依赖静态负载均衡策略，无法感知链路的实时状态（如拥塞或故障）。在智算中心万卡级集群中，此类静态策略可能导致AI训练任务因单路径故障而中断。为解决这一问题，我们提出基于扩展BGP协议的动态路径选择机制，通过链路状态感知实现ECMP路

表3 OLink和Ualink对比分析

内容	Ualink	OLink	说明
生态现状	头部互联网厂商和非英伟达技术体系支持	中兴通讯主导,与部分图形处理器(GPU)客户达成合作意向	中国也有Ualink,阿里集团在积极推动
场景定位	聚焦1K GPU之间Scale-Up互联场景	考虑Scale-Up和Scale-Out融合场景	融合的代价是需要更多地遵循以太网标准,其中OLink支持更大的组网规模
协议架构	按照典型总线型协议设计,基于时钟同步并行传输	基于以太网协议设计,以异步消息通信为主	两者的协议架构思路不同,OLink能够共享以太网生态红利
交换延迟	< 300 ns	< 500 ns	Ualink时延性能优于OLink,但可通过计算和通信重叠消除性能损耗
链路级可靠性	支持链路级重传(LLR)	支持LLR	当前Ualink 1.0版本LLR待完善
原生内存语加载(LD)/存储(ST)	原生支持高并发LD/ST语义	通过Packing方式间接支持LD/ST	Ualink的LD/ST性能优于OLink,OLink可针对特定场景优化性能
直接内存访问(DMA)语义	协议未对DMA进行描述,但技术上支持	原生支持DMA语义	Ualink虽然支持DMA性能,但可能会受线速包长过小的影响
在网计算	1.0版本不支持在网计算,后期版本可支持	支持在网计算	Ualink将在网计算的需求放在1.0版本之后

径的智能负载均衡。

在大规模智算网络场景下,当使用三层CLOS网络架构时,多路径的问题会凸显。然而,传统ECMP技术依赖静态负载均衡策略(如五元组哈希),缺乏对链路实时状态的感知能力,导致流量可能持续转发至拥塞路径,引发AI任务中断。为解决这一问题,我们提出一种基于扩展BGP协议的链路状态感知机制,通过动态通告链路状态信息,实现ECMP路径的智能选择,显著提升网络可靠性和AI任务效率。相较于传统BGP协议仅通告下一跳信息,本方案通过下一跳能力(NHC)扩展TLV(指类型、长度、值格式)字段引入下一跳描述符,使Leaf设备可实时感知Spine-Leaf链路状态(如带宽利用率与故障状态),从而突破第1章所述协议在动态响应能力上的局限。

Spine/Core设备通过在BGP路由通告中引入NHC属性的扩展TLV,通告下一跳和链路描述符,可选携带带宽信息。在链路出现故障或者拥塞(带宽利用率达到一定程度)时,Spine/Core设备将发送控制面或数据面通知消息给Leaf/Spine设备,包括相应的链路描述符以及具体的故障或者拥塞通知。

Leaf/Spine设备在收到Spine/Core设备发过来的协议通告时,建立下一跳、下一跳、链路描述符及链路状态表,以监控下一跳与下一跳之间的链路状态。在收到状态通知消息时,Leaf/Spine设备会更新该表项,并在转发流量时参考该链路状态表决定下一跳设备的选择。为了实现快速的故障恢复,Spine设备在进行故障或者链路带宽预警(预示可能出现拥塞)通告时,通过数据面进行快速通告。需要说明的是,这里的数据面可以基于用户数据报协议(UDP),也可以基于其他类型的IP。Leaf设备可以直接基于链路描述符更新如表1所示的链路状态,以进一步提高处理效率。需要注意的是,当故障链路恢复正常或者拥塞链路不再拥塞时,Spine设备同样需要通告Leaf设备该信息。通告方式可以采用更新属性的方式,以便恢复链路被重新投入使用。

为了避免流量振荡,Leaf设备可以预先向Spine设备发出链路状态订阅消息,即表示自身对某个链路的特别关注。Spine设备在发现该链路出现故障或者拥塞时,直接将报文通知发送给订阅该消息的Leaf设备,实现Leaf设备对该故障的快速反应。这种订阅方式还适用于其他场景,比如:Spine设备在缺省情况下,会在链路带宽占用率达到80%的时候发出带宽告警消息。但某个Leaf设备可能在链路带宽占用率达到50%时就需要得到带宽告警消息。因此,Leaf设备在发送针对链路描述符的订阅信息时,可以携带其要求,比

如要求带宽利用率达到50%时就需要得到通知。Spine设备将记录下这些需求,并在达到这些要求时发送通知消息。

本研究提出的协议创新通过链路状态感知与动态路径选择,解决了智算网络中ECMP路径拥塞与故障导致的AI任务中断问题。实验表明,该机制可降低丢包率90%以上,显著提升大规模智算任务的效率。未来我们将探索与SRv6、Inband Telemetry的深度集成,进一步提升网络智能化水平。

### 3 结束语

智算中心网络作为大模型训练的核心基础设施,其架构设计与技术演进直接影响AI算力的释放效能。本文从拓扑结构、协议优化与组网融合3个维度系统分析了智算网络的演进路径:在拓扑层面,CLOS架构通过分层扩展能力支撑超大规模GPU集群,Dragonfly+/多轨CLOS等变体技术有效平衡了带宽、时延与部署成本;在协议层面,基于扩展BGP的链路状态感知机制,显著提升了ECMP路径选择精度与故障收敛效率。我们在组网方案中提出的OLink融合架构,通过统一以太网协议栈实现Scale-Up与Scale-Out网络的深度整合,突破了传统异构组网瓶颈。实测数据表明,该架构可大幅提升万卡级集群的通信效率。这些技术创新为智算中心构建高带宽、低时延、强扩展性的网络底座提供了实践范例。未来,随着智算中心规模的进一步扩大和技术的不断演进,中兴通讯将持续探索更加高效、灵活的组网技术,以推动智算中心技术的持续发展,为AI和高性能计算的广泛应用提供坚实的技术支撑。

### 参考文献

- [1] 田海东, 张明政, 常锐, 等. 大模型训练技术综述 [J]. 中兴通讯技术, 2024, 30(2): 21-28. DOI: 10.12142/ZTETJ.202402004
- [2] 唐宏, 武娟, 徐晓青, 等. 生成式大模型承载网络架构与关键技术探索 [J]. 中兴通讯技术, 2024, 30(2): 50-55. DOI: 10.12142/ZTETJ.202402008
- [3] REN T Q, LI R P, ZHAO M M, et al. Separate source channel coding is still what you need: an LLM-based rethinking [J]. ZTE communications, 2025, 23(1): 30-44. DOI: 10.12142/ZTECOM.202501005
- [4] 裴丹, 张圣林, 孙永谦, 等. 大语言模型时代的智能运维 [J]. 中兴通讯技术, 2024, 30(2): 56-62. DOI: 10.12142/ZTETJ.202402009
- [5] 韩炳涛, 刘涛. 大模型关键技术与应用 [J]. 中兴通讯技术, 2024, 30(2): 76-88. DOI: 10.12142/ZTETJ.202402012
- [6] 段威, 李和松, 周昆. 智算中心网络技术发展与应用 [J]. 中兴通讯技术, 2024, 30(6): 39-47. DOI: 10.12142/ZTETJ.202406007
- [7] 段晓东, 程伟强, 王瑞雪, 等. 面向新型智算中心的全调度以太网技术 [J]. 中兴通讯技术, 2023, 27(2): 57-63. DOI: 10.12142/ZTETJ.202304011
- [8] QIAN K, XI Y Q, CAO J M, et al. Alibaba HPN: a data center network for large language model training [C]//Proceedings of the ACM SIGCOMM 2024 Conference. ACM, 2024: 691-706. DOI: 10.1145/3651890.3672265

- [9] JIANG Z H, LIN H B, ZHONG Y M, et al. MegaScale: scaling large language model training to more than 10, 000 GPUs [EB/OL]. [2025-03-01]. <https://arxiv.org/abs/2402.15627v1>
- [10] 百度智能云开发者中心. 大规模 AI 高性能网络的设计与实践 [EB/OL]. [2023-06-25]. <https://cloud.baidu.com/article/364290>
- [11] KIM J, DALLY W J, SCOTT S, et al. Technology-driven, highly-scalable dragonfly topology [C]//Proceedings of International Symposium on Computer Architecture. IEEE, 2008: 77-88. DOI: 10.1109/ISCA.2008.19
- [12] SHPINER A, HARAMATY Z, ELIAD S, et al. Dragonfly: low cost topology for scaling datacenters [C]//Proceedings of IEEE 3rd International Workshop on High-Performance Interconnection Networks in the Exascale and Big-Data Era (HiPINEB). IEEE, 2017: 1-8. DOI: 10.1109/HiPINEB.2017.11
- [13] JOUPPI N, KURIAN G, LI S, et al. TPU v4: an optically reconfigurable supercomputer for machine learning with hardware support for embeddings [C]//Proceedings of the 50th Annual International Symposium on Computer Architecture. ACM, 2023: 1-14. DOI: 10.1145/3579371.3589350
- [14] WANG T, SU Z Y, XIA Y, et al. NovaCube: a low latency torus-based network architecture for data centers [C]//Proceedings of IEEE Global Communications Conference. IEEE, 2014: 2252-2257. DOI: 10.1109/GLOCOM.2014.7037143

## 作者简介



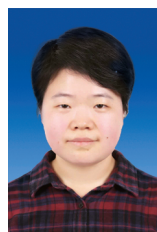
**段威**，中兴通讯股份有限公司资深研发总工，高级工程师；主要从事IP网络、智算中心网络关键技术研究；申请专利30余项。



**于浩**，中兴通讯股份有限公司网络系统专家；主要从事IP网络相关的技术研究和研发工作。



**李和松**，中兴通讯股份有限公司技术规划专家；主要从事智算、光电子以及网络相关的技术研究和规划工作。



**张征**，中兴通讯股份有限公司资深标准工程师；主要从事智算、广域等数通网络相关的技术研究和标准工作。



**王宁诚**，中兴通讯股份有限公司算法专家；主要从事智算网络、智能化相关的技术和算法研究工作。



# 面向5G网络的工业互联网安全技术分析



## Analysis of Industrial Internet Security Technology and Development for 5G Network

陈焱/CHEN Yan<sup>1</sup>, 陈丹/CHEN Dan<sup>1</sup>, 袁琦/YUAN Qi<sup>2</sup>,  
刘小丽/LIU Xiaoli<sup>2</sup>, 徐晓娜/XU Xiaona<sup>2</sup>

(1. 深圳中广核工程设计有限公司, 中国 深圳 518100;

2. 中国信息通信研究院, 中国 北京 100083)

(1. China Nuclear Power Design Co., Ltd. (Shenzhen), Shenzhen  
518100, China;

2. China Academy of Information and Communication Technology, Bei-  
jing 100083, China)

DOI: 10.12142/ZTETJ.202502010

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20240926.1153.004.html>

网络出版日期: 2024-09-27

收稿日期: 2024-07-15

**摘要:** 5G与工业互联网的融合面临接入安全、网络安全、数据安全等诸多安全问题,亟需新的安全技术来应对。总结了面向5G网络的工业互联网安全政策、标准和产业发展情况,研究了面向5G网络的工业互联网在终端、网络、云平台和数据安全方面面临的安全威胁,分析了面向5G网络的工业互联网在终端、网络、云平台和数据方面的安全技术。面向5G的工业互联网安全技术方案,需要立足工业互联网实际应用场景,从终端、网络、云平台和数据安全多个维度引入防护技术,构建全面防护体系,预防安全威胁,提升安全防护水平。

**关键词:** 5G网络安全; 工业互联网安全; 云平台安全

**Abstract:** The integration of 5G and the industrial Internet faces numerous security issues such as access security, network security, and data security, so there is an urgent need for new security technologies to address these issues. This paper summarizes the security policies, standards and industrial development of the 5G-oriented industrial Internet, also studies the security threats faced by the 5G-oriented industrial Internet in terms of terminal, network, cloud platform and data security, and finally analyzes the security technologies of the 5G-oriented industrial Internet in terms of terminal, network, cloud platform and data security. The security technology solution for the 5G-oriented industrial Internet needs to be grounded in the actual application scenarios of the industrial Internet, introduce protection technologies from multiple dimensions including terminals, networks, cloud platforms, and data security, construct a comprehensive protection system, prevent security threats, and enhance the level of security protection.

**Keywords:** 5G network security; industrial Internet security; cloud platform security

**引用格式:** 陈焱, 陈丹, 袁琦, 等. 面向5G网络的工业互联网安全技术分析 [J]. 中兴通讯技术, 2025, 31(2): 72-76. DOI: 10.12142/ZTETJ.202502010

**Citation:** CHEN Y, CHEN D, YUAN Q, et al. Analysis of industrial internet security technology and development for 5G network [J]. ZTE technology journal, 2025, 31(2): 72-76. DOI: 10.12142/ZTETJ.202502010

随着中国5G商用的推进,5G网络建设稳步发展。5G作为新一代移动通信系统,以高带宽、低时延、海量连接等为特性,提升工业互联网的信息化水平,筑牢工业生产的网络基石,为工业的数字化转型注入强劲动力。5G网络的三大场景能够支撑工业互联网的不同功能应用需求,提升网络连接能力和稳定性,推动人、机、物的全面互联,实现工业互联网向智能化、服务化、高端化转型。

5G与工业互联网融合叠加,促进了工业企业数字化、网络化和智能化升级,给工业互联网带来前所未有的发展机

遇<sup>[1]</sup>。但5G与工业互联网的融合也将大量的信息通信技术(ICT)系统威胁和挑战带入工业运营技术(OT)网络中。网络系统的硬件、软件乃至整个系统受到病毒、木马、高级持续性攻击等安全风险的威胁日益加剧,使得面向5G网络的工业互联网在接入安全、网络安全、数据安全和应用安全等方面面临安全威胁。因此,需要提出新的安全技术来应对这些威胁,提升工业互联网的安全保障能力。

目前,面向5G网络的工业互联网安全是业界研究的热点问题。本文将对面向5G网络的工业互联网安全发展情况、

面临的安全威胁以及现行的安全技术进行研究和分析。

## 1 面向5G网络的工业互联网安全发展现状

### 1.1 政策发展现状

国际上,美国与欧盟相继出台了5G和工业互联网相关政策。2016年9月,美国工业互联网联盟(IIC)发布《工业互联网安全框架》,定义了工业互联网的五大安全特性,为工业互联网安全体系建设和部署实施提供最佳实践指南。2019年5月,美国在布拉格5G安全大会上通过了《布拉格5G提案》。2020年1月,美国出台《保障5G安全及其他法案》,同月欧盟发布《5G网络安全欧盟工具箱》。2020年3月,美国发布《美国5G安全国家战略》,该文件正式制定了美国保护5G基础设施的框架,将5G安全进一步提升到更高的战略地位。

中国对面向5G网络的工业互联网安全发展也非常重视,近年来国务院、工业和信息化部等相继出台了各项指导意见和指南等政策,规范并推动5G+工业互联网安全发展。2019年7月,工业和信息化部等十部委联合发布《加强工业互联网安全工作的指导意见》,意见制定了工业物联网安全管理体系,为开展工业互联网安全工作提供了切实可行的指引。工业和信息化部在2019年发布《工业互联网企业网络安全分类分级指南(试行)》,并在2021年1月发布《开展工业互联网企业网络安全分类分级管理试点工作的通知》,以加快构建分类分级管理制度,形成可复制可推广的工业互联网网络安全分类分级管理模式;随后在2021年12月印发《工业和信息化部办公厅关于组织开展工业领域数据安全管理体系试点工作的通知》(工信厅网安函〔2021〕295号),要求企业构建数据安全管理体系,有效保障工业企业数据安全。2022年5月,工业和信息化部又印发了《工业和信息化部办公厅关于开展工业互联网安全深度行活动的通知》(工信厅网安函〔2022〕97号),以健全自主定级、定级核查、安全防护、风险评估等工作机制,提升工业互联网安全保障能力。

### 1.2 标准发展现状

在行业标准方面,2018年5月中国通信标准化协会(CCSA)TC13 WG5开始研制工业互联网安全标准。2021年工业和信息化部发布了数据安全方面的行业标准,并完成了工业互联网平台安全、网络安全等相关标准,目前正在进行数据采集终端设备安全、设备安全防护等标准研制工作。另外TC13 WG5正在进行家电制造行业、钢铁生产、船舶、航

天装备制造和矿山等行业安全应用的标准研制。

在国家标准方面,全国网络安全标准化技术委员会(简称“TC260”)在2018年发布了工业控制系统安全管理、信息安全分级、风险评估等标准,在2019年发布了工业控制系统安全检查指南、工业控制系统漏洞检测等标准,并在2022年发布了工业控制系统信息安全防护能力成熟度模型方面的标准。

随着工业互联网进入规模发展的新阶段,中国需要不断完善工业互联网安全标准,加强面向5G网络的工业互联网安全标准研制工作,提升工业互联网安全防护能力。

### 1.3 产业发展现状

面向5G网络的工业互联网以信息化为基础,以网络化和智能化为核心,成为传统工业企业升级智慧化工厂的核心发展路径。网络安全等级保护2.0进一步扩展了网络安全保护范围,对工业互联网提出了更高的安全扩展要求,有力推动了工业互联网安全产业发展。

面向5G网络的工业互联网安全产业分为安全产品和服务两大类:

1) 工业互联网安全产品分为防护类和管理类产品。防护类产品主要有防火墙、防病毒软件、网络隔离设备、终端/网络入侵检测、工业安全审计等。管理类产品主要有资产管理、补丁管理、身份认证管理、安全运维管理、安全合规管理等产品。防护类和管理类产品成为工业互联网安全整体解决方案的重要组成部分,其市场规模也在持续稳定增长。

2) 工业互联网安全服务分为咨询类、实施类和运营类服务。咨询类主要包括安全评估、安全咨询和安全审计等。实施类主要包括安全集成和安全加固等。运营类主要包括安全应急、安全培训和安全托管等。由于近年来工业网络威胁朝着多样化、复杂化的方向演化,传统的单一安全产品模式已难以满足用户的安全防护需求。在工业互联网安全设计、实施和运维阶段采用全面的工业互联网安全服务,可有力保障工业互联网安全产业发展。

随着中国5G+工业互联网战略的全面实施,政府及企业不断加大安全投入,工业互联网安全产业迎来快速增长期。在政策环境与市场需求的共同作用下,中国工业互联网安全产业进入快速发展的新阶段。

## 2 面向5G网络的工业互联网安全威胁

5G网络在工业互联网的应用,打破了传统工控系统独立、封闭的天然屏障,使得大量工业互联网设备暴露在公共

网络中。虽然很多工业企业对传统网络采取了一定的防护措施,但与系统稳定性、实时性和鲁棒性相比,对网络信息安全性的考虑仍显不足。工业互联网中仍存在大量漏洞和安全威胁<sup>[2-3]</sup>,包括终端安全威胁、网络安全威胁、云平台安全威胁和数据安全威胁,如图1所示。一旦出现安全问题,工业互联网将受到严重影响。

### 2.1 终端安全威胁

5G与工业互联网的融合使得海量工业终端接入成为可能,但同时也给工业互联网带来了更多的攻击风险点。面向5G工业互联网的终端协议种类多样化,能够提供信息量巨大的数据。这些数据分类众多,应用场景多元,但尚无统一的安全标识和认证管理机制,极大增加了接入网络的安全风险。终端设备所用的芯片、操作系统、编码规范、第三方应用软件等,通常存在漏洞、缺陷、后门以及权限滥用等安全问题。海量终端接入面向5G的工业互联网后,一旦被入侵利用,这些终端就有可能形成规模化的设备僵尸网络,成为新的分布式拒绝服务(DDoS)攻击源,进而对工业互联网平台系统和应用等发起DDoS网络攻击。

### 2.2 网络安全威胁

工业互联网网络安全包括5G边缘网络安全和工业企业核心网络安全等。5G网络在边缘计算、能力开放等方面存在安全风险,导致面向5G网络的工业互联网存在网络安全边界不清晰等威胁,给传统基于物理实体隔离的安全防护模式带来极大压力。通过攻击工业互联网的入口,工业互联网中的生产设备在连接到企业核心网络后,非常容易遭受黑客、病毒、高级持续性威胁(APT)、DDoS等攻击。这可能会导致整个网络安全防线失守,进而危害工业生产的稳定运行。

### 2.3 云平台安全威胁

面向5G的工业互联网平台安全主要包括云基础设施安全和云应用平台安全。针对这两个方面的安全威胁主要包括:1)云基础设施有很多虚拟机/宿主机,这些虚拟机/宿主机容易被少数别有用心云租户侵入,通过虚拟机的约束来破坏其他虚拟机或者宿主机;2)不同的云应用设置在不

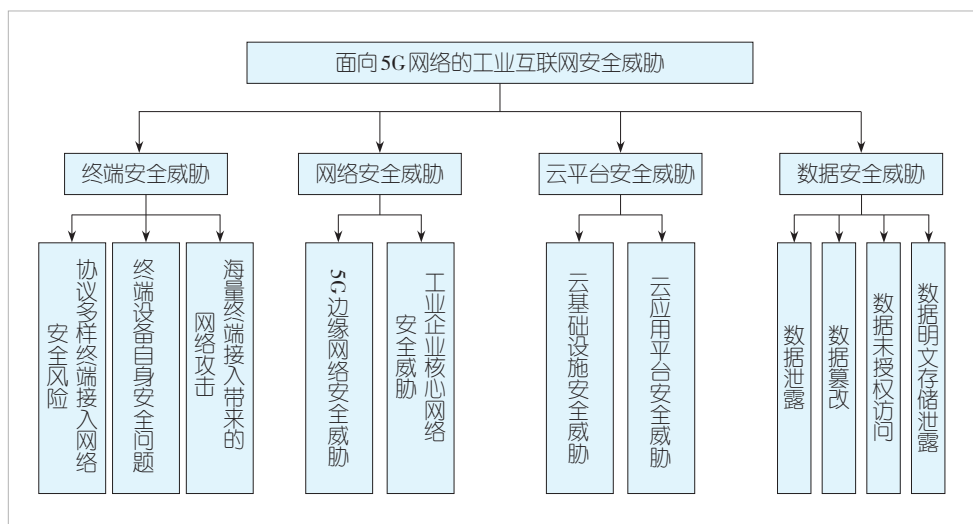


图1 面向5G网络的工业互联网安全威胁

同的服务器上,使得安全边界变得模糊,安全域的划分比较困难,这就可能导致网络和数据安全风险问题;3)云应用平台存在安全漏洞,会使得云服务的安全性得不到有效保障。

### 2.4 数据安全威胁

面向5G网络的工业互联网数据量巨大,且形态多样。在数据采集、传输、存储、处理等各个环节中,存在数据泄露、篡改、未授权访问、明文存储等安全风险。如果无法有效应对这些风险,国家核心行业就会面临安全威胁。这不仅会严重破坏生产的稳定运行,阻碍相关产业发展,还可能对国家网络安全造成冲击,危及国家安全。

## 3 面向5G网络的工业互联网安全技术

5G网络通过提供通信安全、边缘计算安全以及能力开放安全保障,赋能工业互联网安全发展。面向5G网络的工业互联网安全架构,按照信息安全等级保护第三级技术要求进行建设和部署,如图2所示,需要从终端安全、网络安全、云平台安全和数据安全方面来保障自身的安全<sup>[4-5]</sup>,以满足工业企业不同场景的安全需求。

### 3.1 终端安全技术

在终端可信接入安全方面,5G网络对终端接入进行双向认证,在接入网络切片时进行二次认证。在此基础上,还需要建设终端管理系统,使用统一身份认证、国际移动设备识别码(IMEI)绑定、网络准入等复合认证手段,加强接入管控。当发现存在终端设备有风险时,应执行阻断接入操作,将设备下线。



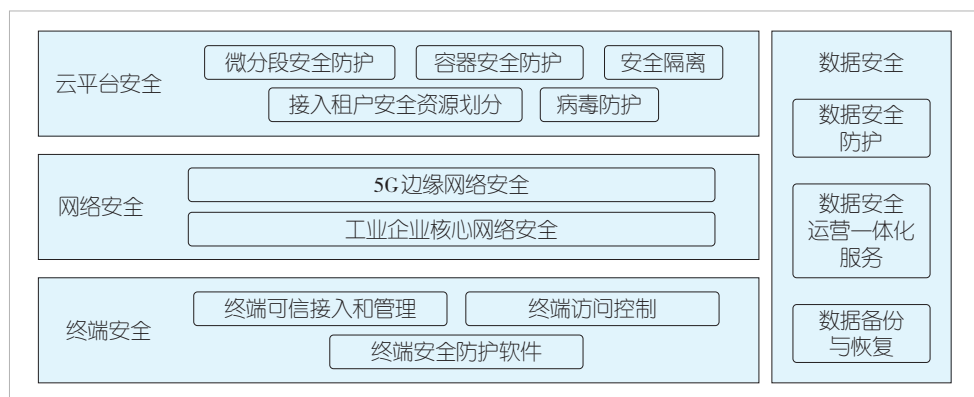


图2 面向5G网络的工业互联网安全架构

通过接入网关加强终端接入控制，并制定网络安全策略（如设置黑白名单的访问控制列表，只允许白名单终端访问），以此实现对接入工业互联网中终端的访问控制。

另外，还需要在终端安装安全防护软件，比如安全防火墙、防病毒软件等，以便及时发现病毒和漏洞，确保接入的终端安全。

## 3.2 网络安全技术

### 3.2.1 5G 边缘网络安全

在面向5G网络的工业互联网中，企业将部署5G边缘计算系统，实现工业互联网业务和应用。具体而言，把原本部署于运营商核心网侧的用户平面功能（UPF），下沉至企业园区内或园区附近，以保障企业业务数据本地存储，提高数据安全性，保证互联网业务和应用的实时性。

在企业园区的UPF侧设置防火墙，可实现UPF侧边界的防护，同时部署入侵防御系统（IPS），可使系统具备入侵防护能力。为了防止边缘计算系统受到DDoS攻击，可以部署全流量检测系统，进行DDoS攻击检测，发现并清洗恶意流量攻击。另外，为了达到安全等级保护三级的要求，还可以部署防病毒网关、堡垒机、主机防护系统等，以增强安全服务能力。

5G边缘计算系统应支持安全通信功能，尽量不使用明文通信。例如：支持互联网协议安全（IPSec）协议，在核心网网元以及边缘计算应用之间建立安全通道，对传输的数据进行机密性和完整性保护，保护传输的数据安全。

5G边缘计算系统应支持网络隔离和访问控制，边缘计算网元的管理、控制、用户平面应分配独立的物理网口，保证多接入边缘计算（MEC）的业务管理、编排控制与用户数据3个平面之间彼此隔离。

5G边缘计算系统可使用OAuth 2.0认证机制对各种访问进行认证和授权，并同时开启自主访问控制和强制访问控

制，遵循最小权限原则，配置合理的访问控制策略，防止非授权访问、篡改等。

### 3.2.2 工业企业核心网络安全技术

为了满足信息安全等级保护第三级的要求，在工业企业网络的核心层旁路部署安全系统，例如防火墙、入侵检测系统、数据库审计系统、堡垒机、漏洞扫描系统等，可满足边界防护、流量分析、日志审计、恶意代码防范

等安全保障要求。

此外，工业企业网络还需要部署高级威胁检测系统。该系统通过对网络行为进行智能分析，实现对网络攻击尤其是新型网络攻击行为的分析，识别其中的已知威胁、未知威胁和异常行为，还原攻击链并进行取证溯源。

针对工业企业网络的业务系统主机和服务器，部署防病毒软件（白名单防护软件）、采用白名单技术，不仅能使主机免受病毒等各种非法攻击，还能有效管控主机的通用串行总线（USB）外部端口。针对Windows系统主机，防病毒软件可以提供适用于工控行业的安全防护，保障关键业务运行，建立稳定运行环境，有效遏制病毒及其变种的运行，满足信息安全等级保护三级的恶意代码防范、安全管理的要求。

## 3.3 云平台安全技术

工业互联网云平台是指基于硬件和软件等资源，为用户提供网络、存储及计算的服务平台。云平台安全通过微分段安全防护、容器安全防护、安全隔离、接入租户安全资源划分和病毒防护实现云平台的安全防护。

1) 微分段安全防护。建立虚拟化环境下的微分段网络安全保护能力，实现虚拟化环境主机层部署微分段安全防护、东西向虚拟下一代防火墙（vNGFW）和虚拟入侵检测与防御（vIDP），以及虚拟机之间的安全保护。

2) 容器安全防护。在运行环境加固方面，利用安全基线对容器主机和容器编排工具进行安全加固，保证容器运行环境安全。在容器网络安全方面，通过部署一定的访问控制安全策略，采用白名单策略实现网络访问控制，实现容器之间、容器与宿主机平台之间、容器与外部网络之间的安全访问控制。

3) 安全隔离。需要在虚拟机之间、虚拟机与宿主机之间进行隔离，也需要对安全边界进行详细的责任划分，对安全域进行进一步细化，对不同的安全域和不同安全级别的安

全域进行不同的访问控制设计。

4) 接入租户安全资源划分。对接入租户进行划分, 利用虚拟专用网络 (VPN)、VRF 和虚拟可扩展局域网 (Vx - LAN) 等技术, 通过在物理网络上构建虚拟的网络层, 为租户提供安全服务。同时为了实现云平台的安全控制与权限隔离, 可通过新一代防火墙、入侵防御系统 (IPS) 和负载均衡产品技术, 为每个租户分配独立非共享的安全防护资源。

5) 病毒防护。通过在虚拟机中部署防病毒软件, 系统可识别并处置针对虚拟机的恶意攻击行为; 同时部署防病毒引擎 (AVE) 网关, 提供病毒查杀功能, 抵御木马、蠕虫、恶意软件的攻击。

### 3.4 数据安全技术

由于面向5G网络的工业互联网场景中会使用海量智能节点、物联网终端设备, 设备和业务应用中包含很多重要数据。因此, 需要针对数据的采集、传输、存储和处理等全生命周期, 采取数据标记、明示用途、数据加密、数据备份、访问控制和安全审计等多种安全防护技术, 建设数据安全运营一体化服务能力, 包括数据资产识别、数据分类分级管理、数据脱敏和加密、数据防泄漏、数据流动监测管理等, 满足复杂化、多样化业务场景下的数据安全治理要求。

与此同时, 在工业企业网络部署数据备份与恢复系统, 不仅能实现实时数据自动备份、异地容灾、热备等功能, 为数据提供可靠性保护, 还能够定期进行备份数据恢复测试, 确保实际备份数据的异机可恢复性, 并进行记录及分析。当业务系统出现故障或数据损坏时, 上述措施能快速恢复数据, 确保重要数据不丢失。

## 4 结束语

面向5G的工业互联网安全技术方案结合工业互联网实际的应用场景, 通过在终端、网络、云平台和数据安全方面引入安全防护技术, 构建全面的安全防护能力, 能够有效预防工业互联网的安全威胁, 提升工业互联网的安全防护水平, 为面向5G网络的工业互联网领域安全发展提供有效的参考。

### 参考文献

- [1] 杨波, 宋翼. 5G 在工业互联网中的安全应用研究 [J]. 通信技术, 2020, 53(11): 2867-2871
- [2] 岳守振, 张琦, 吕金华, 等. 5G+工业互联网网络安全威胁及应对浅析 [J]. 网络安全技术与应用, 2022(10): 100-102

- [3] 张子扬, 韦素萍. 5G 网络中工业互联网安全问题研究 [J]. 中国新通信, 2019, 21(7): 136-137. DOI: 10.3969/j. issn. 1673-4866.2019.07.116

- [4] 尹曾, 王雪源, 吴婷婷. 5G+工业互联网安全防护设计与实践 [J]. 工业信息安全, 2022(10): 76-84

- [5] 范勇杰, 赵磊, 蒋小燕. 5G+工业互联网一体化安全防护体系分析与研究 [J]. 工业信息安全, 2022(10): 85-91

### 作者简介



陈焱, 深圳中广核工程设计有限公司工程师; 主要从事核电、电力工程领域的通信网络业务及信息化、数字化、智能化业务的研究、规划及设计工作。



陈丹, 深圳中广核工程设计有限公司通信主管设计师; 主要从事核电厂无线通信系统和电磁兼容方面的研究工作。



袁琦, 中国信息通信研究院部门主任, 正高级工程师; 主要从事移动互联网、物联网和安全技术领域科研专项、标准等方面的研究工作; 主持或参与国家级或省部级科研专项20余项, 主持或参与制定国家或通信行业标准100余项, 牵头完成重要研究报告80余项, 获省部级奖励9项, 发表论文40余篇, 获授权专利7项, 出版专著3部。



刘小丽, 中国信息通信研究院工程师; 主要从事工业互联网和5G安全技术领域标准、软科学、测试等研究工作。



徐晓娜, 中国信息通信研究院工程师; 主要从事终端安全技术领域标准、软科学、测试等研究工作。

# 物理层安全保密电话的实现



## Realization of Wireline Telephone Based on Physical Layer Security

林立峰/LIN Lifeng, 周子健/ZHOU Zijian, 焦秉立/JIAO Bingli

(北京大学, 中国 北京 800718)  
(Peking University, Beijing 800718, China)

DOI: 10.12142/ZTETJ.202502011

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20240814.1036.002.html>

网络出版日期: 2024-08-15

收稿日期: 2024-07-10

**摘要:** 提出一种基于物理层安全的安全保密电话的实现方法。该方法通过在双向通信的每个信道上释放人工噪声, 使得线路任意窃听点检测到的信号功率远低于噪声功率, 从而无法识别通信信息。合法用户采用同频同时全双工技术消除人工噪声, 恢复正常通信。理论分析与实验结果表明, 该系统能有效抑制自干扰, 使合法用户的信干噪比显著优于窃听者。硬件原型成功研制验证了系统的实用性和有效性, 为高安全性通信场景提供了可行的技术解决方案。

**关键词:** 物理层安全; 同频同时全双工; 保密电话; 人工噪声

**Abstract:** An implementation method for a secure telephone system based on physical layer security is proposed. The security principle involves injecting artificial noise into each bidirectional communication channel, ensuring that the signal power detected by any illegal eavesdropper at any probing point on the line is far lower than the artificial noise power, thereby preventing information extraction. Legitimate users employ co-frequency co-time full-duplex technology to cancel the artificial noise and recover the original communication signal. Theoretical analysis and experimental results demonstrate that this system can effectively eliminate self-interference, allowing the legitimate user's signal-to-interference-plus-noise ratio to significantly outperform that of an eavesdropper, thereby ensuring communication security. The successful implementation of a hardware prototype further validates the system's practicality and effectiveness, offering a feasible technical solution for high-security communication scenarios.

**Keywords:** physical layer security; co-frequency co-time full-duplex; secure telephone; artificial noise

**引用格式:** 林立峰, 周子健, 焦秉立. 物理层安全保密电话的实现 [J]. 中兴通讯技术, 2025, 31(2): 77-82. DOI: 10.12142/ZTETJ.202502011

**Citation:** LIN L F, ZHOU Z J, JIAO B L. Realization of the wireline telephone based on physical layer security [J]. ZTE technology journal, 2025, 31(2): 77-82. DOI: 10.12142/ZTETJ.202502011

随着通信业务种类和规模的快速发展, 信息安全作为隐私保护的核心需求, 已成为现代通信领域的重要组成部分<sup>[1-6]</sup>。在点对点保密通信研究中, 学术界通常采用包含合法用户 (Alice 和 Bob) 与窃听者 (Eve) 的典型模型。相关研究主要围绕两个关键问题展开: 一是确保合法用户之间的安全信息传输; 二是有效防止窃听者的信息截获。

为阐明本文物理层安全方法的特性, 我们先对传统密码学方法进行介绍。该方法起源于两次世界大战期间的密码通信, 其核心机制是将明文转换为密文, 使窃听者 Eve 在无密钥条件下难以还原信息。然而, 此方法存在两点局限性: 1) Alice 需将解码密钥告知 Bob; 2) 即便 Eve 不知晓密钥, 仍可能借助高性能计算设备破解密文。

另一种方法为物理层安全方法, 其理论根基源于 SHANNON 信道容量理论, 后经 WYNER 于 1975 年完善。该

方法以信息比特作为基本单元, 通过数学方法描述信息比特与噪声的关联特性, 从而构建安全容量模型<sup>[7]</sup>。物理层安全的核心目标在于当通信信噪比低于特定阈值时, 即便 Eve 采用任何方法也无法恢复被噪声或干扰所隐匿的信息比特。通常只要 Bob 的信噪比高于 Eve, 即可实现安全通信, 这一结论在安全容量框架下具备坚实的信息论理论支撑。

然而在实际通信环境中, 合法用户难以始终维持对窃听者的信噪比优势。为此, 研究者提出了人工噪声 (AN) 技术来抑制 Eve 的窃听能力。文献[8]指出, 通过在接收机端发射 AN, 利用 Eve 对噪声信息的未知性, 可将其信噪比降低至安全阈值以下。例如, 文献[9]、[10]提出通过部署 AN 成功实现 Alice 与 Bob 间的秘密通信。由于 AN 在接收机 (即 Bob) 处释放, 对于 Bob 而言, 该噪声属于已知干扰, 理论上可完全消除, 从而保障其与 Alice 间的合法通信。近年来, 文献[11]、[12]提出采用无人机搭载干扰器的方案, 旨在借助空间维度拓展, 提升保密通信的灵活性与覆盖范围。

**基金项目:** 国家自然科学基金项目 (62171006)



同频同时全双工 (CCFD) 技术被应用于合法用户的 AN 的消除处理<sup>[13-16]</sup>。该技术利用 CCFD 的自干扰消除机制, 在实现大功率 AN 发射的同时, 可在接收端有效消除 AN 干扰, 从而保障 Alice 通信信号的接收质量。

早期 AN 方法在无线通信应用中取得了一定的效果, 但其存在局限性: 由于无线信号在空中传播时会随距离急剧衰减, 当窃听者 Eve 与合法接收方 Bob 的距离较远时, Eve 接收到的 AN 功率会显著降低, 甚至导致该方法失效。此外, 该方法无法满足信息双向传输场景下的保密通信需求。

针对现有 AN 方法的不足, 本研究基于文献[17]将无线通信场景拓展至有线电话系统。该方法通过电话线注入 AN, 具有两大优势: 首先, 有效克服了路径损耗导致的噪声衰减, 确保所有潜在窃听点的信噪比均低于安全阈值; 其次, 完整保留了电话线路的双向通信功能。下文将详细论述该方法的实现原理、系统架构及实验验证结果。

## 1 工作原理和系统构成

本文中我们提出的物理层安全电话, 其核心思路在于由接收机 (Bob) 向线路发送干扰信号, 使线路上传输的信号被干扰所覆盖。鉴于电话系统具备双向信号传输特性, 为便于阐述基本原理, 我们先对单向信号传输接收的工作机制进行说明, 后续将进一步介绍完整的双向通信保密原理。

### 1) 保密电话工作原理

图 1 (a) 展示了 Alice 向 Bob 发射信号的单向通信保密模型, 两者通过长度为  $L$  的导线进行通信, 窃听者 Eve 在链路中间采用搭线方式进行窃听。将 Alice 发射的通信信号记为  $s_1(t)$ , Bob 在接收该信号的同时发射人工噪声  $n_1(t)$ , 经噪声消除处理后, Bob 可获取  $s_1(t)$  携带的信息; 而 Eve 接收到

的电信号则为  $s_1(t)$  与人工噪声的叠加。

当 Eve 在距离 Alice 点  $x$  处时, 其接收电磁为:

$$r_{\text{Eve}}(x, t) = \sqrt{\alpha(x)} s_1 \left( t - \frac{x}{c} \right) + \sqrt{\alpha(L-x)} n_1 \left( t - \frac{L-x}{c} \right) + n_e(t) \quad (1)$$

在公式 (1) 中  $s_1(t)$ 、 $n_1(t)$  和  $n_e(t)$  分别表示 Alice 发射的信号、AN 和热噪声;  $c$  表示光速,  $\alpha(x)$  描述了电磁波功率沿着导线传播距离的衰减因子。

此时, Bob 在消除 AN 后的接收电磁波可以表达为:

$$r_{\text{Bob}}(t) = \sqrt{\alpha(L)} s_1 \left( t - \frac{L}{c} \right) + \beta n_1(t) + n_b(t) \quad (2)$$

在公式 (2) 中,  $\beta$ 、 $n_b(t)$  分别表示 Bob 处 AN 消除因子和热噪声。强大的消除能力保证了  $\beta \ll 1$ 。

从公式 (1) 和 (2) 可以看出: 一方面, 我们希望 Bob 能够彻底消除 AN, 即满足  $\beta = 0$ ; 另一方面, 需要保证 Eve 处的 AN 功率足够大, 以发挥有效干扰作用。上述约束条件共同构成了本文所提出的 AN 环境下物理安全模型。

### 2) 保密电话系统

我们提出的物理层安全电话为双向通信保密系统, 在电话线两端均配置发射机与接收机, 分别记为通信终端-A 和通信终端-B。其中, 将终端-A 的发射机定义为 Alice-1, 终端-B 的接收机定义为 Bob-1, 对应上述单向保密通信模型; 在反向通信中, 将终端-B 的发射机定义为 Alice-2, 终端-A 的接收机定义为 Bob-2, 由此构建双向通信的两条链路。为避免 Alice-1 与 Bob-1、Alice-2 与 Bob-2 两条通信链路间的相互干扰, 系统采用频率分离技术, 其中 Alice-1 与 Bob-1 链路采用频率  $f_1$ , Alice-2 与 Bob-2 链路采用频率  $f_2$ , Bob-1 和 Bob-2 释放的 AN 亦分别对应  $f_1$  和  $f_2$ 。此外, 通过滤波器实现两个频率信号的隔离。系统结构如图 1 (b) 所示。

因此, 我们仅需对通信链路 Alice-1 与 Bob-1 的物理安全问题展开分析, Alice-2 与 Bob-2 链路的情况与之类似。

### 3) 安全性能分析

安全通信的绝对安全性需在时空维度上同时满足。对于窃听者 Eve 所拦截的信号, 其最大能量可通过对整个信息符号持续时间  $T$  进行积分运算来确定。实际通信中, Eve 可能进行多次拦截, 并在特定时刻实现与 ID 符号同步, 达到最佳检测状态。此外, 还需考虑拦截者在干扰效率最大化时的极端情况。

我们采用矩形波形模型分别表征通信符号与 AN 的相关性。以持续时间  $T$  的首个通信符号为例, 在该时段内, 窃听

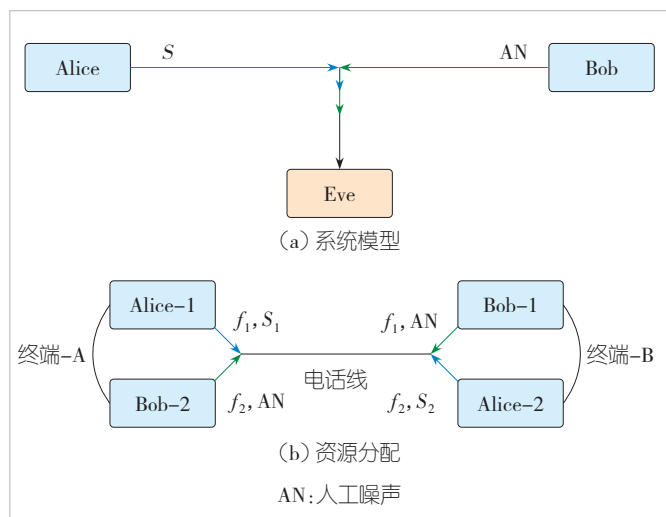


图1 有线安全通信的理论模型

者 Eve 的拦截信号可综合表示为:

$$\hat{r}_{\text{Eve}}^x = \frac{1}{T} \int_0^T r_{\text{Eve}}(x, t) dt = \sqrt{\alpha(x)} s_1(0) + \sqrt{\alpha(L-x)} \hat{n}_1 + \hat{n}_e \quad (3),$$

其中,  $\hat{n}_1 = \frac{1}{T} \int_0^T n_1(t) dt$  和  $\hat{n}_e = \frac{1}{T} \int_0^T n_e(t) dt$ 。同样, Bob 在一个符号持续时间内收到的信号可表示为:

$$\hat{r}_{\text{Bob}}^x = \frac{1}{T} \int_0^T r_{\text{Bob}}(t) dt = \sqrt{\alpha(L)} s_1(0) + \sqrt{\beta} \hat{n}_1 + \hat{n}_b \quad (4),$$

其中,  $\hat{n}_1 = \frac{1}{T} \int_0^T n_1(t) dt$  和  $\hat{n}_b = \frac{1}{T} \int_0^T n_b(t) dt$ 。

为获取最大信干噪比 (SINR), 需综合考虑多次拦截的叠加效应。据此, Eve 拦截信号 SINR 计算方法如下:

$$\gamma_{\text{Eve}}^x = \frac{\alpha(x) \mathbb{E}[|s_1(0)|^2]}{\alpha(L-x) \mathbb{E}[|\hat{n}_1|^2] + \sigma_n^2} = \frac{\alpha(x) P_s}{\alpha(L-x) P_n + \sigma_n^2} \quad (5)。$$

安全通信的最不利情况出现在窃听者 Eve 处于距 Alice 最近且距 Bob 最远的位置。此时, Alice 信号的传播距离达到最大  $L$ , Eve 可获得的 SINR 为:

$$\gamma_{\text{Eve}}^{\max} = \frac{P_s}{\alpha(L) P_n + \sigma_n^2} \quad (6)。$$

相应地, Bob 的 SINR 计算为:

$$\gamma_{\text{Bob}} = \frac{\alpha(L) P_s}{\beta P_n + \sigma_n^2} \quad (7)。$$

我们提出的有线通信系统的容量可以表示为:

$$C_s = [\log_2(1 + \gamma_{\text{Bob}}) - \log_2(1 + \gamma_{\text{Eve}}^{\max})]^+ \quad (8)。$$

容量  $C_s$  表示在窃听者 (Eve) 存在的条件下, 发送方 (Alice) 向接收方 (Bob) 发送可靠且安全信息的最大速率。为了确保绝对安全, 需要满足  $\gamma_{\text{Bob}} > \gamma_{\text{Eve}}^{\max}$ , 即:

$$P_n > \frac{1 - \alpha(L)}{\alpha^2(L) - \beta - \alpha(L)} \sigma_n^2 \quad (9),$$

$$\alpha^2(L) - \beta - \alpha(L) > 0 \quad (10)。$$

由于 AN 消除系数  $\beta \ll 1$ , 且远小于导线衰减因子  $\alpha(L)$ , 因此公式 (10) 在实际应用中通常能够得到满足。需要注意的是, 公式 (9) 和 (10) 构成了保障有线通信系统绝对安全传输的必要条件。导线长度的增加会显著提高系统对 AN 消除能力的要求。

## 2 一种基于 2/4 线电路的 AN 消除方法

我们提出一种基于经典的 2/4 线电路的 AN 消除方法。

AN 消除的原理如图 2 (a) 所示, 其中  $V_s$  表示等效电压源,  $Z_1 \sim Z_6$  是输入端的电阻,  $Z_0$  为输出端的等效阻抗。从 Alice 端看, Bob 作为其负载,  $Z_0$  即为 Bob 电路的等效输入阻抗, 反之亦然。在通信过程中, 一个合法用户向另一个发送  $V_3 \sim V_4$ , 而其接收到的信号是  $V_1 \sim V_2$ 。通过合理设计阻抗参数, 可有效抑制 AN 干扰。由于合法链路的性能主要取决于 AN 的消除能力  $\beta$ , 我们将推导出该系统的自我取消能力  $\beta$ 。

当 Alice 处于信息数据与 AN 发送状态而 Bob 处于空闲状态时 (其电压源未启动), 根据前节分析可知, Alice 接收的信号为  $V_1 \sim V_2$ 。由于 Bob 处于空闲状态且未向 Alice 发送任何信号, 此时 Alice 接收的信号本质上是自身发送的 AN, 因此, Alice 发送的 AN 可通过差分电压信号  $V_1 \sim V_2$  来表示。同理, Bob 发送的 AN 亦可采用相同方法分析, 在此不再重复讨论。

如图 2 (b) 所示, 基于网格电流法, 我们对每个回路应用 Kirchhoff 电压定律, 可得到以下线性方程组:

$$\begin{aligned} (Z_1 + Z_2 + Z_3 + Z_0)I_1 - Z_0I_2 - Z_3I_3 &= 0 \\ -Z_0I_1 + (Z_4 + Z_5 + Z_6 + Z_0)I_2 - (Z_5 + Z_6)I_3 &= 0 \\ -Z_3I_1 - (Z_5 + Z_6)I_2 + (Z_3 + Z_5 + Z_6)I_3 &= -V_s \end{aligned} \quad (11)。$$

上述线性方程组可转化为如下矩阵形式:

$$\mathbf{Z}\mathbf{I} = \mathbf{V} \quad (12),$$

其中,  $\mathbf{Z}$ 、 $\mathbf{I}$  和  $\mathbf{V}$  分别表示系数矩阵、未知数和常数项的列向量。

$$\begin{bmatrix} Z_1 + Z_2 + Z_3 + Z_0 & -Z_0 & -Z_3 \\ -Z_0 & Z_4 + Z_5 + Z_6 + Z_0 & -(Z_5 + Z_6) \\ -Z_3 & -(Z_5 + Z_6) & Z_3 + Z_5 + Z_6 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -V_s \end{bmatrix} \quad (13)。$$

利用 Cramer 法则, 网格电流可以通过公式 (14) 解出:

$$I_i = \frac{\det(\mathbf{Z}_i)}{\det(\mathbf{Z})} \quad (i = 1, 2, 3) \quad (14),$$

其中,  $\mathbf{Z}_i$  表示将  $\mathbf{Z}$  的第  $i$  列替换为  $\mathbf{V}$  后形成的矩阵。矩阵  $\mathbf{Z}_1$ 、

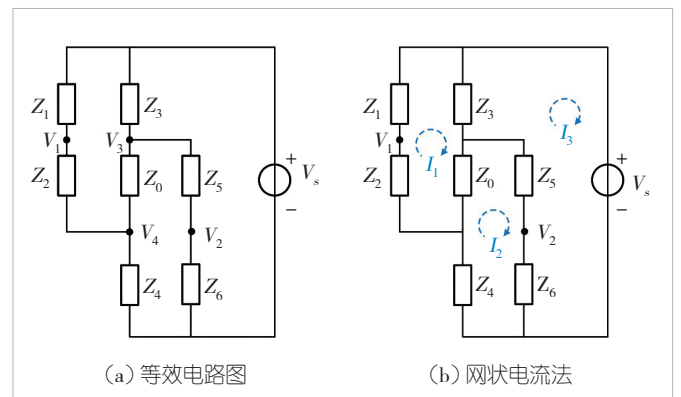


图2 2/4线人工噪声消除电路等效电路图

$Z_2$ 、 $Z_3$ 和 $Z$ 的行列式可以按照方程(15)计算。

$$\begin{aligned} \det(Z_1) &= -V_s \cdot [Z_0(Z_5 + Z_6) + Z_3(Z_4 + Z_5 + Z_6 + Z_0)] \\ \det(Z_2) &= -V_s \cdot [(Z_1 + Z_2 + Z_3 + Z_0)(Z_5 + Z_6) + Z_0 Z_3] \\ \det(Z_3) &= -V_s \cdot [(Z_1 + Z_2 + Z_3 + Z_0)(Z_4 + Z_5 + Z_6 + Z_0) + Z_0^2] \\ \det(Z) &= Z_3 Z_4 (Z_1 + Z_2 + Z_5 + Z_6) + (Z_3 + Z_4)(Z_1 + Z_2)(Z_5 + Z_6) + \\ & Z_0 [(Z_1 + Z_2 + Z_4)Z_3 + Z_4(Z_5 + Z_6) + (Z_5 + Z_6)(Z_1 + Z_2)] \end{aligned} \quad (15)。$$

根据方程  $V_{SI} = V_1 - V_2$  可知, 残余AN过于复杂难以处理。然而, 在实际系统中, 我们可对残余AN进行近似与简化。具体来说, 将电阻值设计为  $Z_1, Z_2, Z_5, Z_6 \gg Z_3, Z_4$ , 可使输出阻抗  $Z_0$  近似等于  $Z_3 + Z_4$ , 即  $Z_0 \approx Z_3 + Z_4$ 。基于此, 可通过舍弃方程中包含小阻值乘积的项, 进一步简化全双工系统中的残余AN信号表达式, 最终得到公式(16):

$$V_{SI} = V_1 - V_2 \approx V_s \left[ \frac{Z_4 Z_5 - Z_0 Z_6}{(Z_3 + Z_4 + Z_0)(Z_5 + Z_6)} + \frac{Z_2 Z_3 + Z_0 Z_2}{(Z_3 + Z_4 + Z_0)(Z_1 + Z_2)} \right] \quad (16)。$$

实际电阻值被设计为  $Z_1 = Z_6$ ,  $Z_2 = Z_5$  和  $Z_3 = Z_4$ 。因  $Z_1, Z_2, Z_5, Z_6 \gg Z_3, Z_4$ , 我们可以进一步假设  $Z_3 = Z_4 \triangleq Z_s$ ,  $Z_2 = Z_5 \triangleq Z_L$  和  $Z_1 = Z_6 \triangleq \alpha Z_L$ 。利用上述符号, 则有  $Z_0 \approx Z_3 + Z_4 = 2Z_s$ 。全双工残余AN可以被简化为公式(17):

$$V_{SI} \approx V_s \left[ \frac{Z_s Z_L - 2\alpha Z_s Z_L}{4(1 + \alpha)Z_s Z_L} + \frac{Z_s Z_L + 2Z_s Z_L}{4(1 + \alpha)Z_s Z_L} \right] = \frac{2 - \alpha}{2 + 2\alpha} V_s \quad (17)。$$

最终, 我们得到AN抑制能力 $\beta$ :

$$\beta = \frac{V_{SI}}{V_s} = \frac{2 - \alpha}{2 + 2\alpha} \quad (18)。$$

图3展示了残余AN信号幅度与公式(18)中参数 $\alpha$ 之间的关系。可以看出, 当 $\alpha = 2$ 时, 系统达到AN抑制效果。然而, 受硬件非理想特性影响,  $\alpha$ 的实际取值往往存在微小误差, 从而降低AN抑制性能。因此, 采用高精度电子元件对优化系统性能具有关键作用。

### 3 性能分析

本章节我们介绍硬件原型的实现方案和实测性能。实验结果表明, 本文所提的保密电话方案具有实用性和有效性。根据文献[17], 信号强度按每100 m距离衰减2 dB的规律递减<sup>[17]</sup>。同时, 设置  $P_s = 10$  dBm,  $\sigma_n^2 = -100$  dBm, 那么发射的AN功率为  $P_n = -60$  dBm。

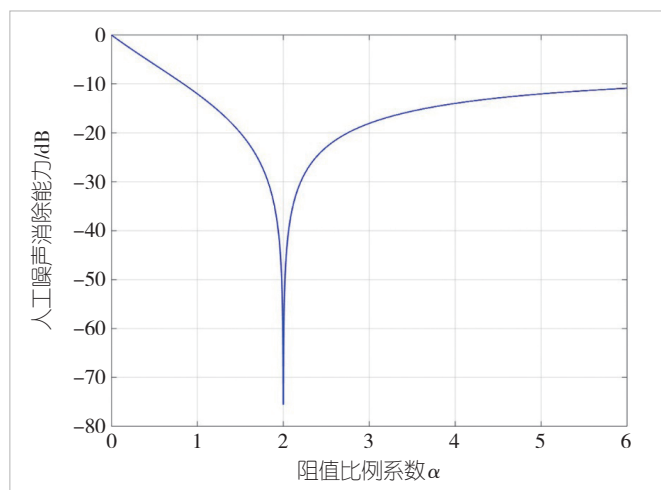


图3 残余人工噪声对应 $\alpha$ 值

硬件系统的实现架构和实物展示分别如图4和图5所示。图4展示了包含Alice和Bob两个通信节点的通信系统架构, 各节点通过电话机与现场可编程门阵列(FPGA)相连, 由FPGA承担信号的发送与接收处理任务。在发送路径中, FPGA将通信信号与AN传输至数模转换器(DAC), 经转换后的模拟信号依次通过放大器与滤波器处理, 再输入至2/4线AN消除器; 接收路径中, 信号经模数转换器(ADC)采样后, 回传至FPGA进行处理。该架构在收发路径中均部署放大器与滤波器, 以保障信号质量与传输性能。音频信号的输入和输出由WM8731执行。基带音频信号处理依托搭载Intel Cyclone V SE 5CSXFC6D6F31C6N芯片的DE10-Standard开发板实现, 基带信号上变频与射频信号下变频分别由AD9767和AD9226芯片执行。系统选用AD8065作为功率放大器, AD8138作为滤波器件, 2/4线电路则基于THS6022与THS6062芯片搭建。

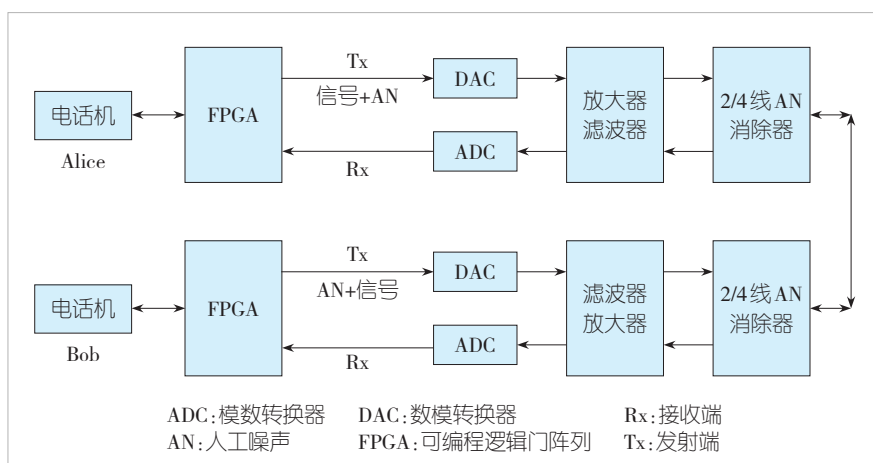


图4 保密电话实现的系统框图

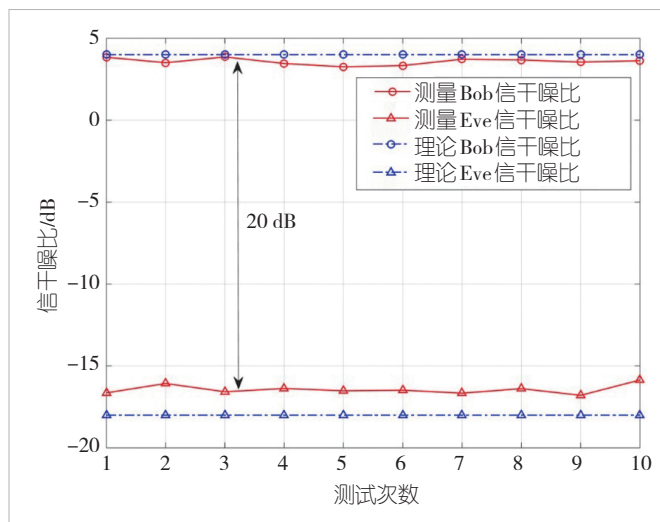
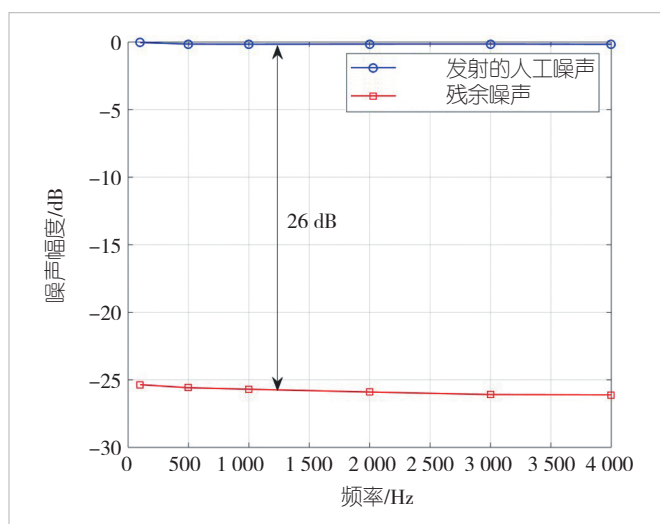




图6展示了信息信号与残余AN的实际频率响应。实验步骤如下：首先，在Alice端生成基带单频正弦波信号；随后，该信号经上变频处理后通过有线信道传输至Bob端。在Alice端，我们通过导线前端采集传输信号，同时在Bob端的导线后端同步接收信号。此外，实验还采集了经2/4线转换处理后的Alice端残余AN信号。通过对接收信号、残余AN信号及传输信号的功率比进行计算，结果表明：该硬件原型具有优异的AN消除性能，可实现至少26 dB的残余AN抑制，同时保持可忽略的信号损耗。

图7对比了合法链路与窃听链路的SINR。实验步骤如下：首先，将1 s时长的音频信号输入合法用户链路，经ADC采样后与AN叠加；随后，叠加信号通过脉冲整形滤波器与功率放大器处理，并经由有线信道传输至另一合法用户。实验结果表明，在所有测试场景中，合法链路的SINR均优于窃听链路约27 dB，与理论值（30 dB）的3 dB偏差验证了所提全双工通信物理层安全方案的有效性与合理性。

图8展示了在26 dB AN消除能力下，系统安全容量与AN功率的关系曲线。安全容量通过公式（8）计算获得，实验测试了100 m、200 m和300 m 3种线缆长度下的性能表



现。结果表明：线缆长度的增加会导致安全容量降低；而提升AN功率则能显著改善系统性能，这源于增强的AN干扰有效抑制了潜在窃听者的信号截获能力。

#### 4 结束语

本文中，我们提出并实现了一种基于全双工技术的新型安全有线通信方案及配套硬件系统。首先，详细阐述了系统设计原理与架构模型，通过信息数据与AN的叠加保障合法用户的通信安全。其次，创新性地提出了一种基于简易电阻网络的AN消除方法，实现了高效的噪声抑制性能。紧接着，重点分析了合法链路的传输性能，其关键性能指标取决于AN消除效果。其中，残余AN水平通过差分信号计算方法精确量化。通过理论仿真与实物实验相结合的方式，充分验证了所提方案的技术可行性与实用价值。

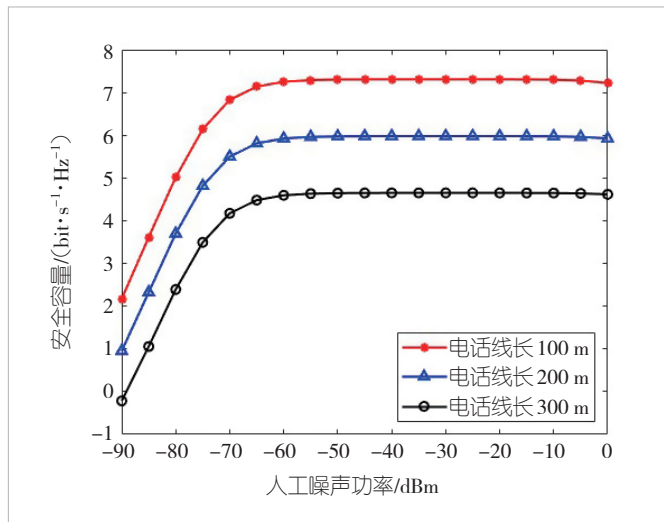


图8 不同电话线长下的安全容量性能对比

## 致谢

烟台大学的姜佩贺老师、湖北民族大学的刘三军老师,以及北京大学的魏来博士对本文的研究做出了贡献,专此致谢!

## 参考文献

- [1] WYNER A D. The wire-tap channel [J]. Bell system technical journal, 1975, 54(8): 1355–1387. DOI: 10.1002/j.1538-7305.1975.tb02040.x
- [2] LIU F, CUI Y H, MASOUIROS C, et al. Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond [J]. IEEE journal on selected areas in communications, 2022, 40(6): 1728–1767
- [3] NGUYEN D C, DING M, PATHIRANA P N, et al. 6G Internet of Things: a comprehensive survey [J]. IEEE Internet of Things journal, 2021, 9(1): 359–383. DOI: 10.1109/jiot.2021.3103320
- [4] CHETTRI L, BERA R. A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems [J]. IEEE Internet of Things journal, 2019, 7(1): 16–32. DOI: 10.1109/jiot.2019.2948888
- [5] SHAFIQUE K, KHAWAJA B A, SABIR F, et al. Internet of Things (IoT) for next-generation smart systems: a review of current challenges, future trends and prospects for emerging 5G-IoT scenarios [J]. IEEE access, 2020, 8: 23022–23040
- [6] QIAO X Q, REN P, NAN G S, et al. Mobile web augmented reality in 5G and beyond: challenges, opportunities, and future directions [J]. China communications, 2019, 16(9): 141–154
- [7] SHANNON C E. Communication theory of secrecy systems [J]. Bell system technical journal, 1949, 28(4): 656–715. DOI: 10.1002/j.1538-7305.1949.tb00928.x
- [8] LI W, GHOGHO M, CHEN B, et al. Secure communication via sending artificial noise by the receiver: outage secrecy capacity/region analysis [J]. IEEE communications letters, 2012, 16(10): 1628–1631. DOI: 10.1109/LCOMM.2012.081612.121344
- [9] SOBERS T V, BASH B A, GUHA S, et al. Covert communication in the presence of an uninformed jammer [J]. IEEE transactions on wireless communications, 2017, 16(9): 6193–6206. DOI: 10.1109/TWC.2017.2720736
- [10] LI K, SOBERS T V, TOWSLEY D, et al. Covert communication in

continuous-time systems in the presence of a jammer [J]. IEEE transactions on wireless communications, 2022, 21(7): 4883–4897. DOI: 10.1109/TWC.2021.3134179

- [11] ZHOU Y, YEOH P L, CHEN H, et al. Improving physical layer security via a UAV friendly jammer for unknown eavesdropper location [J]. IEEE transactions on vehicular technology, 2018, 67(11): 11280–11284. DOI: 10.1109/tvt.2018.2868944
- [12] ZHOU Y, YEOH P L, PAN C H, et al. Caching and UAV friendly jamming for secure communications with active eavesdropping attacks [J]. IEEE transactions on vehicular technology, 2022, 71(10): 11251–11256. DOI: 10.1109/TVT.2022.3186730
- [13] DING J Z, ZHOU Z J, LI W Y, et al. Movable antenna-enabled co-frequency co-time full-duplex wireless communication [EB/OL]. [2025-03-16]. <https://arxiv.org/abs/2401.17049v3>
- [14] DING J Z, ZHOU Z J, WANG C B, et al. Secure full-duplex communication via movable antennas [EB/OL]. [2025-03-15]. <https://arxiv.org/abs/2403.20025v2>
- [15] DING J, ZHOU Z, JIAO B L. New paradigm for secure full-duplex transmission: movable antenna-aided multi-user systems [EB/OL]. [2025-03-15]. <https://arxiv.org/pdf/2407.10393v1>
- [16] MA M, TIAN S Y, CHEN Y Y, et al. A prototype of co-frequency co-time full duplex networking [J]. IEEE wireless communications, 2020, 27(1): 132–139. DOI: 10.1109/mwc.001.1800565
- [17] LIU S J, MA M, LI Y Z, et al. An absolute secure wire-line communication method against wiretapper [J]. IEEE communications letters, 2017, 21(3): 536–539. DOI: 10.1109/LCOMM.2016.2636836

## 作者简介



**林立峰**, 北京大学电子学院在读博士研究生, 正高级工程师; 主要从事同频同时全双工技术的研究工作以及无线收发机研究设计, 研究方向为无线通信、卫星通信、物理层安全; 参与省部级科研项目2项; 已发表论文11篇。



**周子健**, 北京大学电子学院科研助理; 主要从事同频同时全双工技术的研究与相关原理样机的开发工作, 研究方向为无线通信中的信号处理、极化天线系统, 以及软件无线电应用; 已发表论文12篇。



**焦秉立**, 北京大学电子学院教授、中国通信学会智慧医疗专家委员会常务副主任、国家重大专项无限医疗物联网总体规划课题组组长、IEEE 高级会员、北大-时空道宇先进通信联合实验室主任; 主要从事同频同时全双工技术、通信中的信号处理以及移动医疗物联网的研究工作; 先后主持新一代宽带无线通信重大专项、国家“863”课题、国家自然科学基金重点项目等; 已发表论文180余篇。

# 中兴通讯技术杂志社

## 促进产学研合作青年专家委员会

**主 任** 陈 为 (北京交通大学)

**副主任** 秦晓琦 (北京邮电大学) 卢 丹 (中兴通讯股份有限公司)

### 委 员

曹 进 西安电子科技大学

陈 力 中国科学技术大学

陈 为 北京交通大学

陈琪美 武汉大学

陈舒怡 哈尔滨工业大学

陈思衡 上海交通大学

官 科 北京交通大学

韩凯峰 中国信息通信研究院

何 姿 南京理工大学

侯天为 北京交通大学

胡 杰 电子科技大学

黄 晨 紫金山实验室

李 昂 西安交通大学

刘 凡 东南大学

刘春森 复旦大学

刘俊宇 西安电子科技大学

卢 丹 中兴通讯股份有限公司

陆游游 清华大学

宁兆龙 重庆邮电大学

祁 亮 上海交通大学

秦晓琦 北京邮电大学

秦志金 清华大学

史颖欢 南京大学

唐万恺 东南大学

王景璟 北京航空航天大学

王兴刚 华中科技大学

王勇强 天津大学

温森文 华南理工大学

吴泳澎 上海交通大学

武庆庆 上海交通大学

夏文超 南京邮电大学

徐梦炜 北京邮电大学

徐天衡 中国科学院上海高等研究院

杨川川 北京大学

尹海帆 华中科技大学

于季弘 北京理工大学

张 娇 北京邮电大学

张宇超 北京邮电大学

章嘉懿 北京交通大学

赵昱达 浙江大学

赵中原 北京邮电大学

周 伊 西南交通大学

朱秉诚 东南大学

### 刊物相关信息



投稿须知



投稿平台



过刊下载



论文索引与  
引用指南



**办刊宗旨:**

以人为本, 荟萃通信技术领域精英  
迎接挑战, 把握世界通信技术动态  
立即行动, 求解通信发展疑难课题  
励精图治, 促进民族信息产业崛起

**产业顾问:**

段向阳、高 音、胡留军、华新海、刘新阳、  
陆 平、史伟强、屠要峰、王会涛、熊先奎、  
赵亚军、赵志勇、朱晓光

双月刊 1995 年创刊

第 31 卷 总第 182 期

2025 年 4 月 第 2 期

主管: 安徽出版集团有限责任公司

主办: 时代出版传媒股份有限公司

深圳航天广宇工业有限公司

出版: 安徽科学技术出版社

编辑、发行: 中兴通讯技术杂志社

总编辑: 王喜瑜

主编: 王利

执行主编: 黄新明

副主编: 卢丹

编辑部主任: 王萍萍

责任编辑: 徐烨

编辑: 杨广西、朱莉、任溪溪

设计排版: 徐莹

发行: 王萍萍

编务: 王坤

《中兴通讯技术》编辑部

地址: 合肥市金寨路 329 号凯旋大厦 1201 室

邮编: 230061

网址: tech.zte.com.cn

投稿平台: tech.zte.com.cn/submission

电子信箱: magazine@zte.com.cn

电话: (0551) 65533356

发行方式: 自办发行

印刷: 合肥添彩包装有限公司

出版日期: 2025 年 4 月 25 日

中国标准连续出版物号: ISSN 1009-6868  
CN 34-1228/TN

定价: 每册 20.00 元