

CAICT 中国信通院



# 智能算网 (AI Fabric 2.0) 研究报告

2025年9月

# 目录



## 编制说明

01

## 01 前言

02

## 02 数据中心网络发展趋势与挑战

03

2.1 数据中心网络产业发展趋势 ..... 03

2.2 通算数据中心网络发展趋势与挑战 ..... 05

2.3 智算数据中心网络发展趋势与挑战 ..... 06

## 03 数据中心网络代际演进

09

## 04 AI Fabric 2.0 关键技术

11

4.1 AI 网元 ..... 11

4.1.1 绿色超宽 ..... 11

4.1.2 高速光互联 ..... 12

4.1.3 内生安全 ..... 15

4.1.4 智能遥测 ..... 16

4.1.5 AI交换机 ..... 17

4.2 AI 联接 ..... 18

4.2.1 新型网络架构 ..... 18

4.2.1.1 多平面组网技术 ..... 18

4.2.1.2 对等组网技术 ..... 20

4.2.2 网络级负载均衡技术 ..... 23

4.2.2.1 流级负载均衡技术 ..... 23

4.2.2.2 逐包负载均衡技术 ..... 24

# 目录

4.2.3	拥塞控制技术	26
4.2.3.1	端网协同拥塞控制技术	26
4.2.3.2	FlexBuffer拥塞控制技术	28
4.2.3.3	拉远训练精准流控技术	29
4.2.3.4	AI ECN 2.0技术	30
4.2.4	在网计算技术	31
4.2.5	在网存储技术	31
4.2.6	高稳韧性技术	33
4.2.6.1	故障恢复技术	33
4.2.6.2	闪启技术	36
4.2.6.3	光链路检测技术	37
4.2.6.3	网络互联冗余	39
4.3	AI大脑	41
4.3.1	仿真验证/孪生仿真	41
4.3.2	自动化Agent	42
4.3.3	故障Agent	42
4.3.4	网维Copilot	43
05	总结和展望	45



# 编制说明

本研究报告在撰写过程中得到了多家单位的大力支持，在此特别感谢以下顾问专家、参编单位和参编人员：

顾问专家（排名不分先后）：

中国信息通信研究院云计算与大数据研究所所长 何宝宏

华为数据通信产品线数据中心网络领域总裁 王武伟

参编单位（排名不分先后）：

华为技术有限公司、中国信息通信研究院（云大所）

参编人员（排名不分先后）：

郭亮、王少鹏、王建兵、钱骁、李建高、张白、李军、韩涛、吴圣伟、陈太尚、魏智杰、耿煜、屈飞园、赵笑可、李久勇、李杰、汪若虚、管紫轩、侯延祥、温华锋、陈龙、焦雪松、许建、吴洋、胡秀丽、赵科学、张力、李晨飞

# 1 前言

产业数字化的快速发展推动了数据中心的演进，数据中心网络的发展经历了三个主要阶段：虚拟化阶段、云化应用阶段和算力服务化阶段。

在虚拟化阶段和云化应用阶段，数据中心为办公和生产系统提供虚拟化和云化基础设施，数据中心网络采用传统以太网技术用于支撑数据的集中管理以及计算和存储资源的池化应用。发展到算力服务化阶段，随着高性能存储、超算中心高性能互联和AI算网的引入，数据中心网络需要提供更高带宽、更低时延以及更高可靠性。然而，传统以太网技术无法满足要求，导致数据中心高性能存储选择FC专网承载，超算中心高性能互联则通过IB专网承载，而AI算网则通过增强的以太网技术承载。这种三种网络技术并存大大增加了数据中心网络的管理运维复杂度和成本。

与此同时，计算技术的广泛应用和数据中心对以太网交换需求的显著增长，推动了以太技术的快速发展。随着400GE以太技术的商用和800GE以太标准的发布，以太网技术不仅在带宽上超越了IB网络，无损以太网技术也逐渐成熟，为高性能网络向无损以太演进奠定了基础。

超融合以太以实现数据中心网络融合为目标，将通用计算、存储、高性能计算统一承载在0丢包以太网技术栈上，实现从三张网到一张网的融合部署，统一网络架构，推动无损网络向超融合网络架构演进，实现算网融合。

在当前数字化浪潮席卷全球、AI应用呈指数级快速增长的时代背景下，数据中心网络作为算力承载与数据流通的关键枢纽，其性能优劣直接决定了企业数字化转型与智能化创新的成败。基于对这一行业痛点的深刻洞察，以及推动AI产业生态良性发展的使命感，我们精心筹备并推出这份《智能算网（AI Fabric 2.0）研究报告》（下面简称AI Fabric 2.0）。

相对于传统数据中心网络，AI Fabric 2.0不仅是技术上的迭代升级，更是从理念到架构、从技术到应用的全面革新，是对重新定义了AI时代数据中心网络的价值范式。



# 数据中心网络发展趋势与挑战

## 2.1 数据中心网络产业发展趋势

随着数字经济的加速发展，人工智能、大数据、云计算、区块链等新一代信息技术已成为社会生产力的关键支撑，尤其是以GPT-4、Llama和DeepSeek为代表的大语言模型（Large Language Models, LLMs）在人工智能领域取得突破性进展，引发了对算力资源前所未有的需求。

在数字经济加速发展的背景下，算力已成为新型基础设施的重要支撑，是推动经济社会高质量发展的核心驱动力。作为数字经济的关键组成和引领性领域，算力正加快成为培育新质生产力、赋能产业升级的重要引擎。全球范围内，各国围绕算力展开了激烈竞争，纷纷通过政策强化算力产业的引导，以抢占数字经济发展的制高点。

例如，美国、欧盟等均已将算力网络纳入国家数字基础设施建设规划。2019年11月，美国发布《国家战略性计算计划》，提出开发、拓展并提升国家的计算基础设施和生态系统。2020年9月，欧盟对“欧洲高性能计算共同计划”进行了升级，重点发展下一代超级计算，强化欧洲数字主权。2023年，欧盟在《2030年数字十年政策方案》也提出“加强欧盟范围内传输、计算和数据基础设施建设”，以实现欧盟2030年数字化转型的共同目标。此外，沙特制定了《沙特愿景2030》战略，旨在发展高性能计算和人工智能技术，提高在全球竞争力。

我国也在持续强化算力产业政策引导，统筹通用算力、智能算力和超级算力，为网络强国和数字中国建设筑牢数字基座。2021年3月，“十四五”规划首次提出“全国一体化大数据中心体系”，建设E级与10E级超算中心；同年5月，国家发改委等四部委联合发布《全国一体化大数据中心协同创新体系算力枢纽实施方案》，明确提出布局全国算力网络枢纽节点，启动实施“东数西算”工程。2021年7月，工业和信息化部发布《新型数据中心发展三年行动计划（2021-2023）》，提出构建以新型数据中心为核心的智能算力生态体系，构建具备高技术、高算力、高效能、高安全特征的新型算力基础设施。2023年10月六部委印发《算力基础设施高质量发展行动计划》（工信部联通信〔2023〕180号），提出2025年算力规模超过300EFLOPS（每秒浮点运算次数），智能算力占比35%的目标，

并阐述了算力内网络高性能传输调度要求，以及无损网络技术应用要求。数据中心网络连接算力，其性能直接决定了整个数据中心的实际算力水平。以大模型训练为例，需要同时协调数千张甚至数万张算力卡资源，数据吞吐量成为AI计算的关键瓶颈，需要高效的数据流水线支持，对网络的带宽、时延和可靠性都提出了极高要求。因此，要提升数据中心算力服务能力，就必须进一步提升数据中心网络性能。根据《中国综合算力指数（2024年）》报告，计算、存储和网络是算力最重要的组成部分，在数字经济发展新时期，需要更加关注构建集群算力、存力和运力于一体的数据中心。

## 2.2 通算数据中心网络发展趋势与挑战

根据中国人民银行发布的金融行业标准《金融数据中心容灾建设指引》（JR/T 0264—2024），通用计算网络容灾正在从“备份可用”向“智能韧性”升级。这一指引不仅提升了对灾难恢复能力，更是金融等行业在通用计算核心业务与数据中心发展理念上的重大转折。在央行《指引》的推动下，金融行业的容灾体系正在加速变革，而政务、泛互联网等领域的通算核心业务也在共同推动数据中心网络进入颠覆性重构期。

未来十年，数据中心的核​​心将发生根本性变化，从“异地容灾预防灾难”向“极端容灾预防战争”演进。以国内某大型银行为例，其业务架构从“业务主备、数据灾备”向“业务多活、数据多活”持续演进，这将推动通算数据中心服务器数量增长10倍，单个数据中心向“十万级”集约化建设迈进。预计在未来10年内，国内大行将率先在内蒙古、西安、成渝等地分阶段建设大型数据中心，最终建成50万甚至百万服务器级别的“多地多中心”数据中心架构，这一趋势将给数据中心的网络架构和运维带来新的变化和挑战。

### 1) 核心演变：从“孤岛保障”走向全局智慧

传统数据中心网络是维系单一计算资源池运作的静态管道，其主要功能是保障计算与存储之间的基本连接与稳定运行。然而，伴随业务多活、数据多活（双活/多活）架构的深入应用，网络角色正重构为连接全国乃至全球分布式计算资源的“神经中枢”。这一中枢需实时掌握所有业务组件的分布状态，并依据负载、灾害风险与延迟感知自主调整流量路径。它不再仅仅是通用计算资源的连接者，而是智能调度各资源池承载能力的核心大脑，确保客户各类核心业务在任意位置发生，都能由最优计算节点响应处理。这就要求异地数据中心间具备全局流量调度能力，可根据链路延迟变化动态切换支付结算类关键业务路径，实现用户“零感知”切换。

### 2) 超大规模与智能韧性协同演进挑战加剧

伴随通用计算服务器迈向50万乃至百万级规模，单数据中心网络的物理规模与逻辑复杂性激增。传统依赖人工策略管理的模式已无法满足需求。带宽对等化、毫秒级切换等硬性要求，迫使网络架构向深度智能化演进：

**意图驱动网络（IDN）与AI融合：** AI的应用将网络运维从故障后的辅助诊断扩展到运行风险预测和优化。运维系统将基于对业务意图的理解（如“支付交易必须在50毫秒内返回结果”）和实时网络状态结合，自主计算最优路径，并自动执行调整，无需人工干预。同时，借助AI技术，在网络变更或故障处理时，能够实现智能化处置，真正迈向“无人值守”数据中心。

**性能极限与新协议普及：** 多活架构下，数据中心间海量数据同步是关键性能瓶颈。低延迟、高带宽的RDMA（如RoCEv2）将普遍应用于金融核心交易与关键数据库同步环境，对无损以太网能力提出严峻挑战。此外，传统Spine-Leaf的CLOS架构是否适用超大规模的数据中心，如何应对多POD间的大规模流量灵活调度，也将面临新的挑战。

**“战争级”韧性催生容灾与加密升级：** 为应对地质灾害及冲突破坏，如何支撑数据中心网络“中枢”在极端环境下的通信能力成为关键瓶颈。网络需要基于智能故障感知恢复，助力网络常稳业务永续。同时，为保障跨境字、跨DC此类高速链路互联场景的传输安全，通信安全等相关技术也将加速在高韧性DC架构中落地。

未来十年数据中心网络将彻底超越传统连接的定位，真正成为驱动金融、政府等业务韧性、智能与效能变革的根基性力量。其成功演进绝非单纯的技术升级，更将成为行业在数字经济时代制胜的关键能力基石。谁能率先完成网络从“业务保障通道”向“智能韧性运营枢纽”的跃迁，谁就将在未来激烈的数字化和智能化竞争中占据先机。

## 2.3 智算数据中心网络发展趋势与挑战

AI产业正迎来前所未有的快速发展期，呈现出几大显著趋势变化。

1) **大模型规模指数级增长，头部OTT大模型持续摸高，大模型训练带动网络发展模型性能竞赛白热化：** 2025年上半年，OpenAI的GPT-4.5、Google的Gemini 2.5 Pro、Anthropic的Claude 4等模型在复杂推断能力上持续突破。例如，Gemini 2.5 Pro在数学推理（2025年高考数学卷得分140分）和多模态理解（支持100万令牌上下文窗口）上达到新高度，而Claude 4 Opus通过混合推断模式实现“快速响应”与“深度反思”的动态平衡。

**模型摸高驱动集群规模走向10万卡，代际快速演进：** 北美头部互联网客户在智算领域的投资保持60%增速，例如Meta在2025年持续扩大英伟达H100 GPU集群，规模超35万卡，网络代际向800GE快速演进；“星际之门”总投资超1150亿美元，2025年底建设超

30万卡GB200集群。国内客户也在积极跟进，例如字节2025年在智算领域投资1600亿元人民币，持续AGI（通用人工智能）研发；阿里发布Qwen2.5-Max，未来3年投资3800亿投入AI和云基础设施。

模型尺寸从最初的数十亿参数发展到如今的万亿参数级别，带动算力规模呈现爆炸式增长，相应的网络规模和带宽需求也随之急剧增长。这直接推动了网络速率从Gb级别迈向400GE/800GE/1.6TE高速时代，对网络延迟、吞吐量和可扩展性带来了严峻挑战。为了应对这些挑战，网络架构必须进行根本性变革，传统的多层网络结构由于转发延迟高、配置复杂等问题已难以满足AI流量的需求，扁平化的网络架构成为必然选择。

## 2) 推理业务快速崛起，并呈现多样化发展，带动网络发展

随着大模型技术逐渐成熟，智算行业正经历从“训练优先”向“推理主导”的战略转型。这一趋势由两大核心驱动力：

**算力结构重构：**2025年中国智能算力规模预计达1037.3 EFLOPS（是通用算力的12倍），随着模型的成熟以及生成式人工智能应用的不断拓展，推理场景的需求日益增加，推理服务器的占比将显著提高。IDC数据显示，预计到2028年，推理工作负载占比将达到73%。

**应用场景爆发式增长：**智算业务的应用场景正从互联网行业向金融、医疗、制造、交通等传统行业全面延伸，推动各行业的智能化转型。例如在金融领域，智算技术被广泛应用于风险控制、智能投顾、欺诈检测等场景。通过对海量金融数据的分析和建模，智算中心能够为金融机构提供更精准的风险评估和投资建议，提高金融服务的效率 and 安全性。

AI的应用场景从单一的训练模式扩展到丰富的推理服务，包括实时推理、批量推理、边缘推理等多种形态，业务流量从规律化转向复杂化和多样化。与训练场景相对单一的流量模式不同，推理业务呈现出显著的异构性和动态性特征。业务调度流量、集合通信流量、KV Cache访问流量、存储访问流量等多种流量类型在同一网络中并存，且流量模式具有明显的突发性和不可预测性。同时，推理部署形态也发生革新，PD分离，Attn-FFN分离等新架构模式相继涌现，有效优化资源利用效率，进一步推动推理集群化部署。

## 3) 普惠AI与终端革命：应用AI化，要求高性能、低时延的网络

AI技术正从“奢侈品”变为“基础设施”，带动产业AI升级，并深度融入终端设备。

产业AI升级，对性能要求更高：以金融为例，风控策略从之前的静态模型+动态策略升级为基于AI模型的实时主动防护，要求在交易的同时快速做出判断。以某国内领先的银行为例，其日均交易额约1.03万亿元，日均交易笔数达数亿级别，每秒交易上万次，会产生大量的高频并发，需要超千卡的大规模、高性能的推理网络。

**AI应用升级，对时延更加敏感：**AI应用从当前的对话类/推荐类业务，将向AI助手/智能体/云服务演进，随着应用的升级，例如远程AI医疗、量化交易等要求<5ms的时延，带来网络时延降低200倍的诉求。

**智算行业正经历从“算力堆砌”到“效率革命”的质变：**训练向推理的演进重构了商业价值链条，头部模型的技术突破定义了行业标杆，普惠AI与终端革命则打开了万亿级市场空间。未来，技术融合、生态协同与绿色转型将成为竞争焦点，企业需在“技术领先性”与“场景落地性”之间找到平衡点，方能在这场算力革命中占据先机。

#### 4) Agentic AI带来通用计算和AI计算混合部署，网络归一化探索

头部企业通过参数规模扩张和网络架构革新保持技术领先，同时探索标准化网络方案以应对算力碎片化挑战。

客户探索通智合一，统一网络架构的路线：快手在2025年互联网大会，提出“一盒打天下”的理念，不同业务通过归一化网络承载，复用一套硬件设备，通过单一架构实现最优性价比（可降低30%建网成本）

随着Agentic AI的发展，通用计算和AI计算将混合部署，传统的计算网络、存储网络 and 智能网络正在加速融合，形成统一的智算网络基础设施。在这一融合进程中，以太网凭借其开放生态、成本优势和持续的技术演进能力，正成为业界的主流选择。相比之下，InfiniBand虽然在智算内部高带宽互联方面具有优势，但其封闭生态和高昂成本限制了广泛应用，面对推理集群中多种流量混合场景以及未来推理异构卡部署需求显得力不从心。现如今，业界产生了诸如UEC、SUE、ETHLink等多种基于以太网设计的协议范式，展现出更好的开放性和适应性，这预示着以太网在AI产业中的重要地位将进一步提升。

# 3 数据中心网络代际演进

在数字经济向智能世界加速迈进的进程中，算力已成为核心生产力引擎，而网络作为算力流通的“血管”，其效率与韧性直接决定数字价值释放。随着AI技术的蓬勃发展以及云化架构的不断升级，数据中心网络正站在变革的前沿。

AI Fabric 2.0以“算力满载、业务永续”为价值目标，构建“AI网元-AI联接-AI大脑”三层架构，重塑算力网络底层逻辑，为智能时代的算网协同与业务创新提供基石支撑。



图1 AI Fabric 2.0三层架构

## AI网元：硬件级算力传输的“坚实底座”

作为AI Fabric 2.0的硬件基石，由交换机和光模块构成，以四大核心能力为AI业务筑牢算力传输根基。针对AI训练“大带宽、突发性”与推理“高并发、低延迟”的差异化需求，它通过FLB均衡、Token均衡算法实现95%以上吞吐效率，400ns级低时延（远优于业界750ns水平）精准匹配AI实时性诉求；在安全层面，构建“态势感知和多级加密”防护体系，保障高价值训练数据与模型参数0泄露；同时内置AI芯片，具备ns级高精度统计与故障可视化能力，让网络异常可感知、故障可追溯，从物理层确保算力传输高效、安全、可控。

## **AI联接：协同级算力调度的“高效枢纽”**

AI联接打破传统网络被动传输的局限，以“端网协同、全网可靠、智能调度”重塑算力流转路径，成为打通算力供需的关键中枢。通过端网协同与在网计算技术，实现端侧与网络的深度联动；依托磐石高可靠架构的iReliable技术，结合光模块AI能力，可实现毫秒级故障切换，网络可靠性提升10倍以上，确保业务7×24小时不中断；同时，通过训练网络级负载均衡、推理调度算法，动态优化算力分配，避免节点过载或闲置，推动AI训练和推理性能整体提升10%以上，让算力从“静态分布”转向“动态高效流动”。

## **AI大脑：智能级网络管理的“指挥核心”**

作为AI Fabric 2.0的顶层管控中枢，AI大脑以“网络、业务、运维”三级自动化，彻底重构网络管理模式。在网络层面，通过网安一体架构与AI策略算法，实现网络与安全配置的统一分析、自动生成，变更准确率达100%，杜绝人工配置错误；在业务层面，凭借Rubook拖拽式编排工具与10+预置场景API，无需代码开发即可完成流程对接，将业务上线周期从数月压缩至周级；在运维层面，构建应网一体体系，不仅能主动感知应用质量、一键定位端到端路径故障，实现分钟级排障，还能结合知识图谱与AI算法，提前发现5大类40多种高频风险隐患，覆盖70%场景，让网络管理从“被动应对”转向“主动预判”。

AI Fabric 2.0为数据中心网络发展指明了方向，也成为应对未来超大规模算力集群高性能算力需求的重要解决方案。以实现数据中心网络融合充分释放算力为目标，实现算力网络融合。

# 4 AI Fabric 2.0关键技术

## 4.1 AI 网元

### 4.1.1 绿色超宽

随着AI、机器学习、5G、物联网等新兴技术的发展，数据中心内部和之间的数据流量呈爆炸式增长，对网络带宽提出了更高的要求。为了满足这一需求，数据中心网络服务器的带宽升级（端口速率从10GE升级到400GE，Serdes从10G到112G）。与此同时，对网络芯片的吞吐诉求也随之上升，需要更大容量和规格芯片。

通算网络市场，又分为EDC市场和OTT市场。EDC市场主要包括金融、政企、运营商等市场，对带宽演进的诉求相对来说慢于OTT市场，一般5年左右更新代际，当前处在从10GE接入到25GE接入演进的节奏周期中。接入设备存在框盒架构和盒盒架构两种，一般采用10GE/25GE接入、40GE/100GE汇聚设备，近几年10GE和40GE的需求量在逐步下降，25GE/100GE逐步上升。

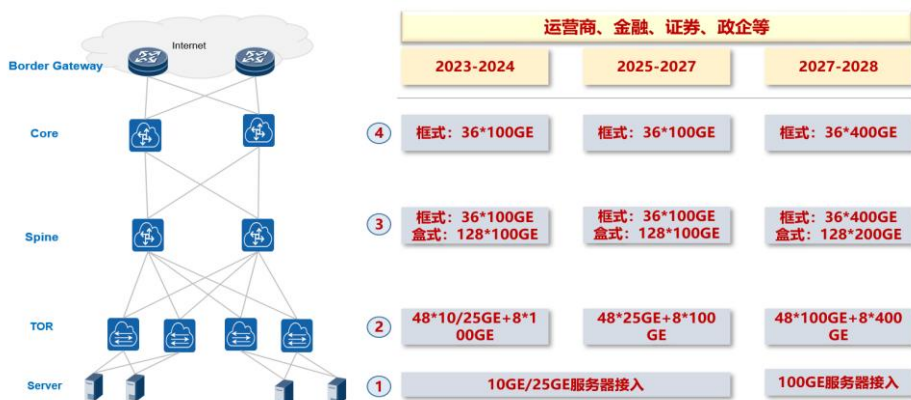


图2 EDC市场对带宽演进诉求

对于OTT市场，带宽演进诉求比较强烈，当前处于从100GE服务器向200GE服务器过渡的阶段，预计26年开始200GE服务器代际部署，采用24\*400GE+8\*800GE或48\*200GE+8\*800GE作为接入，128\*400GE或128\*800GE作为汇聚。

智算网络市场当前则以400GE接入为主，采用盒盒组网或框盒组网，采用比如32\*400GE盒式、128\*400GE盒式、36\*400GE框式等款型；后续会向800GE演进，比如采用128\*800GE等设备进行组网部署。两层盒盒的扁平化组网仍是最优的选择，两层盒式51.2TE的盒子通过光shuffle或者多芯片拼接盒子即可构建多平面的10万卡的集群，满足大规模的训练组网的诉求。

随着设备带宽进一步增大，设备本身的散热和机房散热也成为客户关注的焦点。盒式设备功耗从2KW~3KW上升到5KW~6KW，框式设备功耗增加到40~50KW，已经逐渐逼近风冷散热的极限，客户机房部署也需要更多的空调来解决散热问题，机架利用率下降，同时担心出现局部热点。因此，液冷设备逐步进入到行业视野范围，部分客户已经开始计划部署液冷交换机，降低PUE，提升机架利用率，解决高功耗设备部署问题。

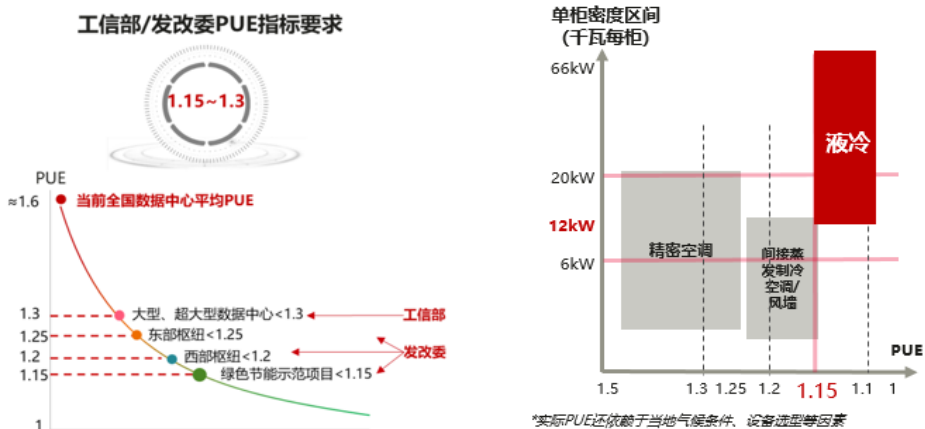


图3 国家多PUE指标要求

### 4.1.2 高速光互联

随着AI网元朝着高速、高密、大集群的方向演进，传统电互连受限于高频损耗成倍提升，传输距离随速率提升受限，无法满足集群发展的要求，“光进铜退”的互联演进趋势不断凸显。另一方面，超高速率光模块对网元设备的功耗压力不断提升，诸如LRO和NPO的新型低功耗互联方案，也是未来光电技术演进的重要方向。

Serdes代际	DAC电互连传输距离	Optical光互连传输距离
56G	3m	10m到2km/10km
112G	2m	10m到2km/10km
224G	1m	10m到2km/10km

表1 “光进铜退”的光电互联趋势

参考IEEE 802.3系列以太网物理层标准，随着Serdes速率提升，无源铜缆的传输距离从56G代际的3m下降为224G代际的1m，除此之外为了实现多通道高速率，并行无源铜缆也带来了不易弯折不易部署的困难。

与电互联相比，光互联的传输距离得益于光纤的低损耗特性，传输距离不随基带信号的提升而衰减。从56G到224G代际均可通过PAM4光信号调制、直接探测方案和光数字信号处理（oDSP）技术的可插拔光模块，可以轻松实现A数据中心所需的100m内短距服务器接入和2-10km交换设备互联。依赖并行光纤、波分复用技术等光学技术，可以轻松通过轻质化易部署光缆实现单端口400GE/800GE和1.6TE的数据传输。

因此，综合传输距离、可演进性和可部署型多方面因素，光互联解决方案会逐渐替代电互联下沉入柜。



图4 400G QSFP-DD FR4和800G OSFP SR8可插拔光模块和DLC光缆

### 新架构光互联方案

随着端口速率的提升，传统光模块的单体功耗不断增大，与此同时单交换机端口数量也在不断提升，两者叠加对整机散热和功耗带来了巨大的挑战。近些年，一些新架构光互联方案不断被提出。基于现有的可插拔形态，在当前主流的收发端均有DSP的可插拔光模块基础上，分别衍生出了仅具备发端DSP，收端算法集成在主芯片上的LRO（Linear Receive Optics）方案和收发端DSP均集成在主芯片上的LPO（Linear Pluggable Optics）方案。更进一步的，在牺牲了可插拔方式的易维护性，将光器件/光引擎从模块内部转移到PCB板或主芯片基板上的过渡形态。

被称之为NPO(Near-Packaged Optics)和CPO ( Co-Packaged Optics )，其获得了因电互联距离降低带来的功耗收益和相比于长链路的LPO稍强的可互通性。而被称之为光互联的“终极形态”的方案是OIO (Optical Input-Output) 芯片出光架构，该架构完全去除了当前电Serdes的串并转换，交换芯片IO直接驱动高密光引擎，变成真正意义上的光电混合芯片。

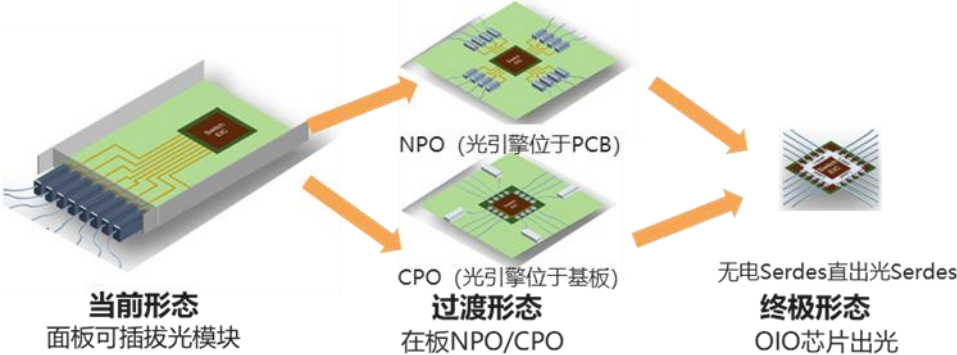


图5 光互联方案架构演进

以下表2从功耗、可演进性、可维护性和可互通性几个维度评估了几种新架构的光互联方案。从对比中可以发现，可插拔光模块创新架构中,LRO具备更好的综合优势，非可插拔架构中，NPO-OIO的技术路线可能是未来AI网元光互联演进的可行路径。

方案	可插拔光模块			NPO	CPO	OIO
	DSP模块	LPO模块	LRO模块			
功耗	高, 18pj/bit	低, 10pj/bit	中, 12pj/bit	低, 9pj/bit	低, 6pj/bit	低, 4pj/bit
可演进性	高, 448G	低, 224G存疑	中, 224G+	中, 224G+	中, 224G+	高, 448G+
可维护性	高, 易运维	低, 难运维	中, 可运维	低, 难运维	低, 难运维	低, 难运维
可互通性	高	低	中	中	中	低

表2 多维度对比新架构优劣

# 4.1.3 内生安全

数据中心一般承载高价值的业务，容易成为黑客攻击的目标；随着业务的上云会增加网络的边界，对加密传输的诉求在增加；网络设备安全容易被忽视，成为薄弱环节，形成攻击的跳板。数据中心网络设备需要有更高的安全防护。

应对日益增加的安全攻击，需要设备具备原生的安全能力：

1. 出厂可信：开放透明、安全无后门；
2. 开局安全：客户证书+华为签名确保安全入网；
3. 管理安全：国密加密，保证了数据的安全传输，日志独立存储空间，安全防篡改；
4. 协议安全：OSPF、BGP等动态路由协议支持国密SM3算法，路由邻居可信；
5. 链路安全：光模块PhySec和链路MACsec国密加密；
6. 主动感知：安全配置主动核查审计；
7. 安全底座：硬件可信根防止软件、文件被篡改，完善的安全算法防止运行性代码注入。



图6 设备原生安全能力

通过下列的措施可以提供全生命周期的安全保护，避免出现安全事件，受到监管机构的处罚。

1. **硬件可信根：**软件、补丁、特征库、特性包完整性校验，防篡改；分层密钥组件，根密钥在CPU中保护，难破解；
2. **关键应用国密加密：**SSH、路由协议、MACsec全面支持国密算法；
3. **代码级安全防护：**完善的安全函数、安全编译技术，防止运行性代码注入漏洞。

## 4.1.4 智能遥测

随着人工智能迈入大规模时代，其赖以生存的“神经网络”——AI智算网络，正以前所未有的速度和复杂度急剧膨胀。数以万计的GPU/NPU集群通过高速互联构成庞大的计算实体，处理着PB级的数据洪流。然而，这种规模的扩张也带来了前所未有的“黑盒”挑战：错综复杂的网络链路、瞬息万变的流量模式、以及分布式训练中同步与通信的微妙状态，都变得难以用传统工具洞悉。因此，对网络可视化的诉求已从“锦上添花”演变为“不可或缺的生命线”。我们不再仅仅满足于看到拓扑图和流量仪表盘，而是迫切需要一种深度、智能且具预测性的全景可视化。

随着DCN的规模的扩大，网络设备需要提供**快速的故障感知和运维的能力**。

**黑盒监控**：iFIT，随流检测，实现丢包和时延的网络级监控。

**白盒上报**：PacketEvent，针对丢包的报文进行聚合上报。

**业务可视**：Service Telemetry，实现RDMA业务的网络时延测量和IO可视。

**智能流量分析**：实现对指定的业务流进行深度分析，得到指定业务流的丢包率，时延（纳秒级）等高精度性能指标数据，并支持将分析结果输出至分析器进一步展示和分析，从而监控网络运行情况和快速定位网络故障。

另外网络设备的可靠性和运维的能力不能解决端侧故障的场景，通过端网协同可以实现E2E的能力，比如华为数据面故障穿越技术和iFIT的端网协同就可以实现整个数据中心网络的故障越过和随流检测的能力。

**全流分析**：一种网络流量监控和分析技术，针对TCP/UDP/VXLAN/RoCEv2流量有效。

全面监控整网流量，以便及时检测和分析异常，对网络运维至关重要。传统网络流量监控方式各有特点，但均无法监控整网流量信息。

**NetStream**：是一种基于接口的统计流量，鉴于CPU压力，很难实现1:1采样，不能精确反映流量状态。

**基于ACL的流量统计**：可以精确统计流信息，但需要提前配置ACL。一般用于故障发生后的故障定位，无法实时感知故障。

**镜像**：可以将报文镜像到分析器集中处理，但镜像报文主要为控制面报文，缺乏对转发面报文的检测（如TCP的SYN/FIN报文）。镜像占用ACL资源，如果进行全流镜像，又会造成分析器压力过大。

全流分析对进入到设备的流量根据五元组字段建立流表，进行流量统计和流量异常检测，

并能上报流表信息进而完成网络质量分析。全流分析技术可以实现对全网流量的分析，它借助于设备内置芯片，支持1:1采样，不影响转发性能。同时提供关键事件上报能力，减轻上层分析软件的处理负担。

另外网络设备的可靠性和运维的能力不能解决端侧故障的场景，通过端网协同可以实现E2E的能力，比如华为数据面故障穿越技术和iFIT的端网协同就可以实现整个数据中心网络的故障越过和随流检测的能力。

全流分析：一种网络流量监控和分析技术，针对TCP/UDP/VLAN/RoCEv2流量有效。

全面监控整网流量，以便及时检测和分析异常，对网络运维至关重要。传统网络流量监控方式各有特点，但均无法监控整网流量信息。

NetStream：是一种基于接口的统计流量，鉴于CPU压力，很难实现1:1采样，不能精确反映流量状态。

基于ACL的流量统计：可以精确统计流信息，但需要提前配置ACL。一般用于故障发生后的故障定位，无法实时感知故障。

镜像：可以将报文镜像到分析器集中处理，但镜像报文主要为控制面报文，缺乏对转发面报文的检测（如TCP的SYN/FIN报文）。镜像占用ACL资源，如果进行全流镜像，又会造成分析器压力过大。

全流分析对进入到设备的流量根据五元组字段建立流表，进行流量统计和流量异常检测，并能上报流表信息进而完成网络质量分析。全流分析技术可以实现对全网流量的分析，它借助于设备内置芯片，支持1:1采样，不影响转发性能。同时提供关键事件上报能力，减轻上层分析软件的处理负担。

## 4.1.5 AI交换机

AI交换机通过内置AI引擎和AI芯片，将AI能力融入软件架构和操作系统，形成智能主控板、智能转发芯片和智能光模块等核心AI部件。与传统交换机相比，AI交换机将AI能力嵌入设备的软硬件和操作系统，设备通过集成AI芯片增加设备智能面，通过设备控制面、管理面、智能面的相互协同，实现AI流量识别、AI加速、AI查表转发、AI节能、AI安全防护等能力，实现算力满载，业务永续。

AI芯片是AI交换机的最核心的智能处理单元，采用深度学习识别大小流、感知时延和吞吐、预测趋势，可支持流量精准控制和精细化调度。AI主控板是AI交换机的“大脑”，

通过集成通用算力、智能算力和AI算法，将AI嵌入软件架构和操作系统，构建交换机的高性能计算平台，为网络数据传输和控制提供智能化的支撑。智能光模块可以实时检测光模块亚健康故障，结合网络智能体精准定位故障点，大幅提升网络运维效率。

## 4.2 AI 联接

### 4.2.1 新型网络架构

#### 4.2.1.1 多平面组网技术

网络层级增多将显著提升建网成本和降低网络设备综合效率，层级越高将有更多的端口资源用于网络间互联而非直连算力，因此降低算力集群的组网层级是未来创新技术方向之一。

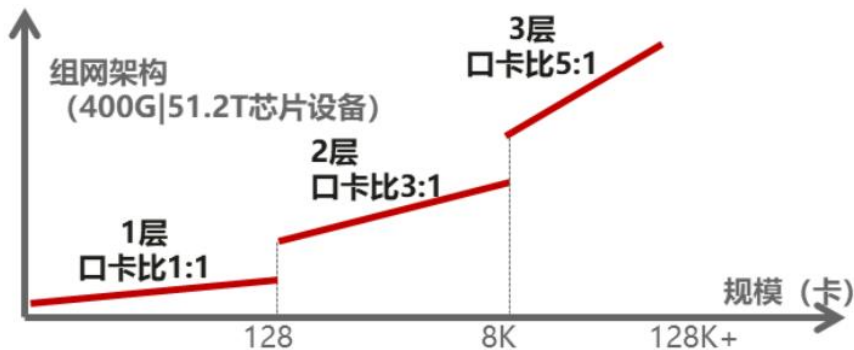


图7 特定交换芯片容量模型下算力规模与组网架构的关系

如上图所示，不同层次组网中网络端口数与算力卡端口数的比例分别是：

一层组网，网络端口数同算力卡端口数是1:1。

两层组网，Leaf无收敛情况下，网络端口数同算力卡端口数是3:1。

三层组网，Leaf/Spine无收敛情况下，网络端口数同算力卡端口数是5:1。

多平面组网技术是指在算力集群规模不变的情况下，将一组完全互联的Clos多级组网，拆分成多个彼此平行独立的低层级Clos组网，以实现扁平化网络架构。

当前51.2T高性能网络成为主流。51.2T交换容量的设备普遍为112G Serdes设计，可以通过512\*100GE Radix扇出对外提供更多的连接端口。使用100GE端口形态的两层Clos架构可以构建128K（512\*256）个100GE接入带宽的网络。

当算力卡支持将400GE带宽性能拆分为4\*100GE端口，分别接入4个100GE接入带宽的Clos网络，组成4平面扁平化组网架构。如下图所示，红，橙，蓝，绿分别组成一个独立的Clos网络。

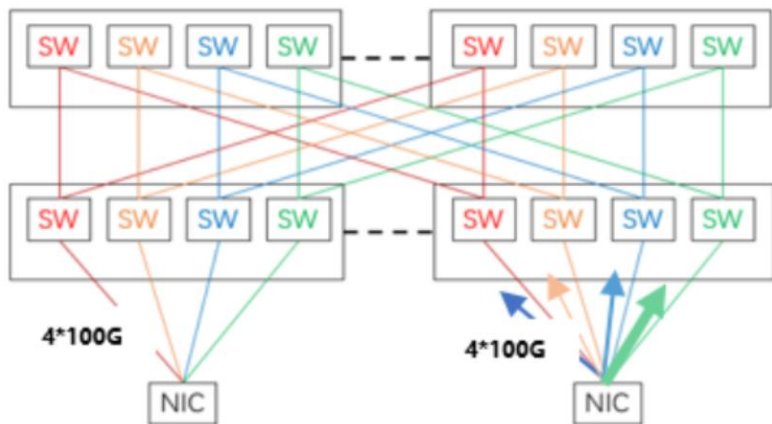


图8 四平面扁平化组网架构

通过多平面扁平化组网技术可以实现两层Clos建超万卡规模集群。如下图所示以512\*100GE扇出为例，多平面两层组网可以将口卡比为3:1的集群规模由8K大幅提升16倍到128K区间。

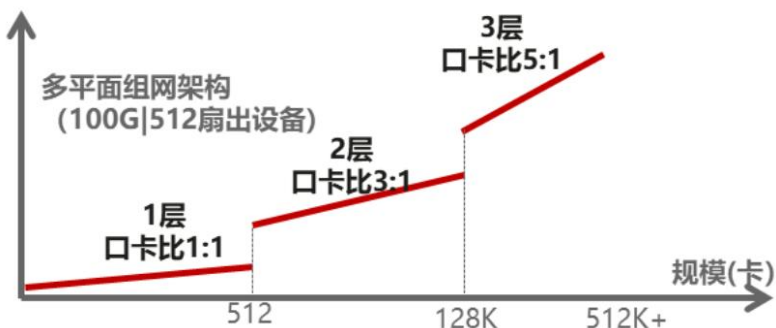


图9 多平面扁平化组网规模

## 4.2.1.2 对等组网技术

2020年12月国家发改委高技司发布《关于加快构建全国一体化大数据中心协同创新体系的指导意见》为契机，牵引全国大体量数据中心向八大枢纽布局。围绕国家重大区域发展战略，在京津冀、长三角、粤港澳大湾区、成渝，以及贵州、内蒙古、甘肃、宁夏等地布局建设全国一体化算力网络国家枢纽节点，发展数据中心集群，引导数据中心集约化、规模化、绿色化发展。

关系国计民生行业如金融银行，运营商等，响应国家一体化大数据中心规划，纷纷在中西部布局其下一代数据中心。企业依据过往20年数据中心发展，结合当前及未来自身业务发展规划，均考虑建设数十万服务器体量的大园区数据中心，园区由4~20不等的单元楼构成。

以金融大型银行为例，当前运行的数据中心通常部署在1~3单元楼内，每栋楼总服务器在2~3W量级，因此使用延续了很多年的集中式Clos组网架构。集中式Clos架构在一定规模&楼宇布局的背景下具备简单、稳定、易部署易扩容的属性。核心层一般为框式可插卡交换机设备，框式设备随CloudRegion建设节奏扩容核心层单板即可满足扩容诉求，且网络设施和算力设施在同一L0故障域范围，只需考虑供电，机房故障场景即可提供高可用组网。

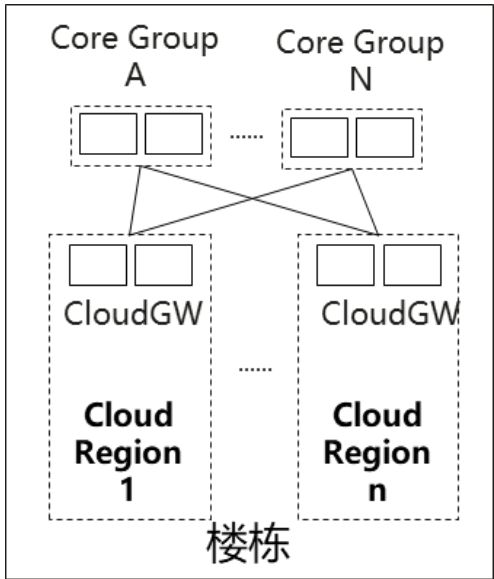


图10 当前金融大型银行数据中心组网逻辑架构

未来随着大型多楼宇数据中心建设，集中式纵向增长的网络架构不能很好适应业务发展，主要问题在于故障半径增加，端到端网络跳数和转发时延增加，数据中心网络需要架构性变革适应业务变化。

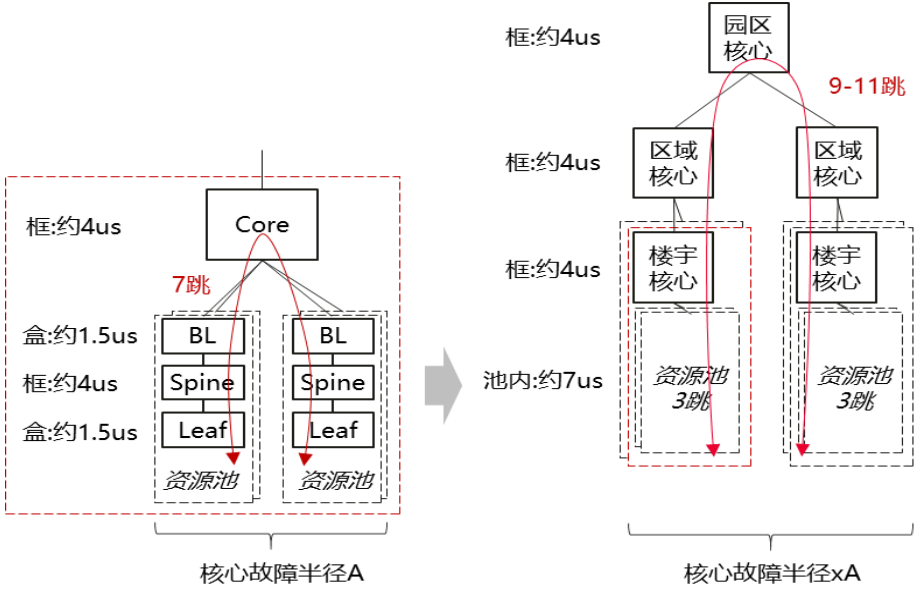


图11 业务规模变化对集中式组网的挑战

如图11所示，原数据中心单楼宇资源池+核心端到端7跳，网络时延约18us，核心故障半径为单楼栋3~5W服务器体量。

使用集中式纵向扩容方式建设数十万服务器体量的园区数据中心，涉及增加多楼宇间区域核心（基于规模可选），多区域间再汇聚到园区核心，多层汇聚后端到端7跳增加到11跳，时延由18us增加到34us增幅88%，顶层核心故障的影响半径涉及全园区范围，劣化数倍。针对上述挑战，金融大规模园区数据中心去中心化架构势在必行，网络由纵向扩展向横向扩展变化。分布式去中心化业务单元间对等架构更适用于未来发展。

分布式去中心化业务单元间对等架构，也叫做对等组网架构，主体设计思路是减少纵向汇聚层级，在某一层面业务单元间横向互联，业务单元可以是云为单位，也可以以机房为单位或者以楼宇单位设计，下面我们以楼宇为单位展开讨论。

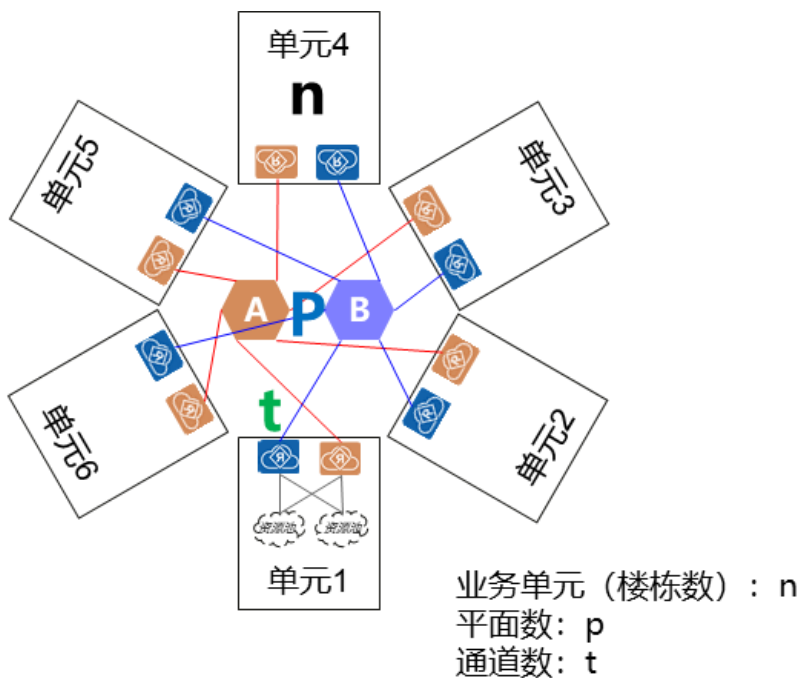


图12 对等架构示意图

如上图12所示，6个业务单元间的对等架构考量，数据中心整体去中心化结构，每个业务单元彼此对等。基于数据中心整体业务容量及可用性规划，业务单元之间建议使用多平面、多通道组网技术满足诉求。

在对等架构下任意业务单元失效，仅影响该业务单元本身，对存量其他单元无影响，故障爆炸半径缩小为集中式的 $1/n$ 。任意两个业务单元间E2E跳数增加1跳，可支撑的组网规模扩大十余倍满足未来中长期业务扩容演进，时延由 $18\mu s$ 增加到 $22\mu s$ ，较集中式组网优化 $40\% \sim 60\%$ 。除组网架构本身带来的比较优势外，业务单元间流量控制也有很多创新工作待开展。例如单元间直连最短路径和非最短路径的选路控制问题，上图任意业务单元间最短路径容量为 $p \cdot t$ ，非最短路径容量为 $p \cdot t \cdot (n-1)$ ，两类容量均为多业务单元间所有互访流量共享，因此需要全局视角的路由算法适配，以满足不同业务等级的要求。

截至本书成文，对等架构已经在部分行业客户试点，随着创新的深入，有望成为下一代大规模数据中心网络的关键技术之一。

## 4.2.2 网络级负载均衡技术

AI集群训练过程中，参数通过高速互连网络在不同的服务器间进行同步交互，这些通信流量具有共同的特征：流量大且成周期性、流数量少，且并行任务间有强实时同步性要求，通信效率取决于最慢的节点。上述的流量特征导致网络较易出现负载分担不均、整网吞吐下降的问题，从而影响AI集群训练的性能。

当前网络均衡的主流技术有两种：**流级负载均衡**和**逐包负载均衡**（有些文献中也叫包喷洒，Packet spray）。

### 4.2.2.1 流级负载均衡技术

基于ECMP的流级负载均衡是当前DCN默认采用的算法，该算法通过流的五元组进行哈希选路，在流数量较多的场景下，基本能实现整网均衡；但ECMP在流数量较少时存在哈希冲突问题，导致算力损失。现在业界多厂家持续在流级负载均衡做技术创新，引入全局视野，整体规划所有业务流量，从拥塞后调整到事前整网调度流量，实现整网无冲突，成为业界主流。

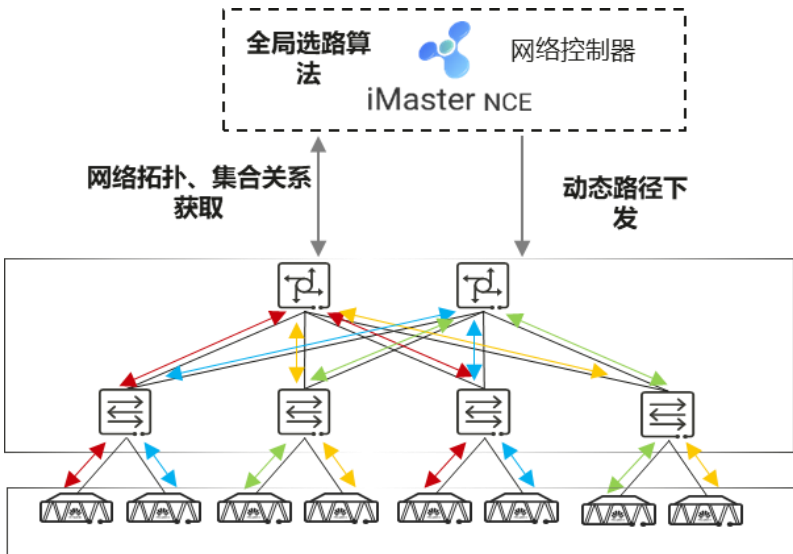


图13 华为网络级负载均衡原理图

例如华为提出网络级负载均衡，通过算网协同，网络集中控制器从计算侧获得整网的流量矩阵，将全局流量划分到独立通信域，为每个通信域类的业务流计算全局最优路径，确保整网流量无冲突，在训练场景中效果如下：

	华为网络级负载均衡	ECMP	提升
HCCL allreduce	38.5GB/s	24GB/s	60%
Bloom模型（样本数/s）	59.5	50.7	17%
VGG16模型吞吐（图片数/s）	13,998	11,394	23%
LLaMA模型吞吐（样本数/s）	22.5	17.1	32%

表3 华为网络级负载均衡与传统ECMP对比

除训练业务外，AI推理业务同样面临流量冲突问题。与训练场景相比，推理业务的流量具有更高的动态性和突发性，传统的全局调度技术难以有效适应其变化特征。针对这一挑战，华为基于推理业务的实际特点，自主研发了AI推理调度算法。该算法有效提升了系统的调度效率，实现了首token时延降低37%、推理并发度提升12.5%，显著增强了推理系统的整体性能与响应能力。

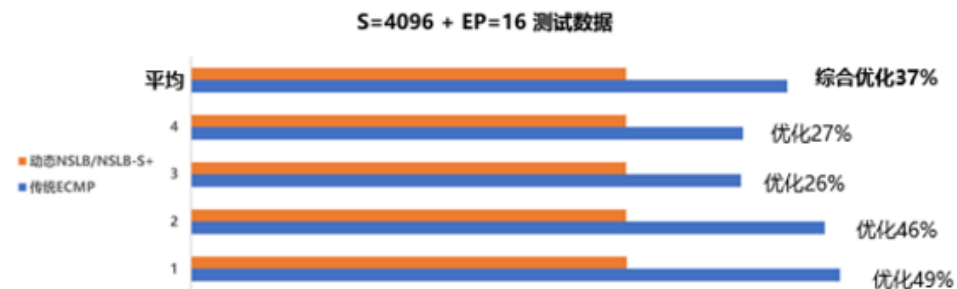


图14 推理算法与传统ECMP对比

#### 4.2.2.2 逐包负载均衡技术

逐包负载均衡（Per-Packet Load Balancing）的核心机制在于：基于数据包粒度动态分配流量，交换机依据报文到达顺序，将流量均匀调度至所有可用链路。该机制在理想条件下可实现链路带宽的100%利用率，保障网络整体吞吐性能最大化。

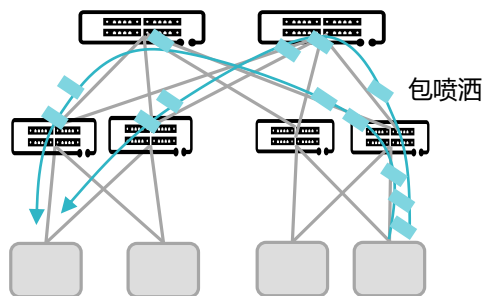


图15 逐包负载均衡技术原理图

然而，在无损RoCE网络场景中，逐包负载均衡面临两大关键性技术挑战：PFC扩散风险和包流混合干扰，这使得传统逐包均衡技术难以直接应用于高性能无损网络。逐包负载均衡需解决如下两个问题：

### 问题1：网络中PFC反压扩散

PFC ( Priority-based Flow Control, 优先级流量控制) 作为保障RDMA网络无损传输的核心机制，已在大规模RDMA网络中广泛部署。在智算集群场景中，慢节点（例如受CPU、内存、PCIe或网卡性能瓶颈影响的节点）会触发PFC反压并向网络扩散。尤其在逐包负载均衡模式下，由于数据包被均匀分散至所有路径，PFC反压帧也将发送至全网链路导致全网级死锁。

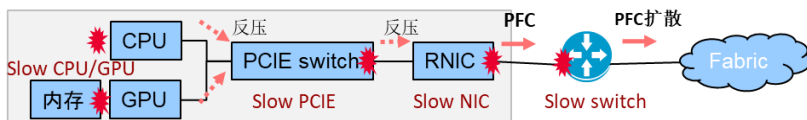


图16 PFC反压扩散

为解决此问题，需将PFC反压严格约束于第一跳接入交换机（即目的Top-of-Rack交换机，DTOR），阻断其向核心网络的扩散，从而彻底消除对非相关发送端的性能干扰，显著提升网络整体吞吐效率与可靠性。

### 问题2：包级均衡和流级均衡业务混合部署

在云环境中，不同租户可依据需求指定网络采用包级或流级负载均衡策略，导致网络中包级流量与流级流量并存。

包级流量天然均匀分布于所有路径，维持整体均衡；而流级流量则集中于特定路径（或部分端口）。由于承载流级流量的路径带宽需求显著高于其他路径，极易引发局部拥塞，从而破坏包级均衡的全局性。

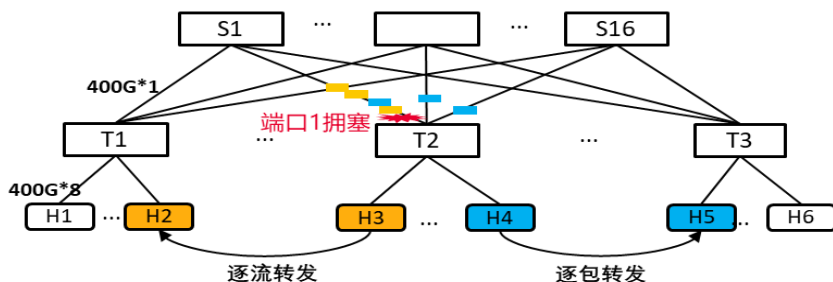


图18 包级和流级均衡业务混合部署

为解决此问题，需引入增强型包流混合负载均衡技术。其核心思路在于动态感知拥塞：当检测到拥塞路径时，主动降低包级流量在该路径上的转发比例。实测表明，相较于传统包级负载均衡基线方案，本技术可实现整网吞吐量提升47%。

## 4.2.3 拥塞控制技术

### 4.2.3.1 端网协同拥塞控制技术

端网协同拥塞控制技术采用“端（智能网卡）网（交换机）”配合的方式达到交换机上的近似“零排队”时延，使得端到端传输时延接近静态时延。

早期端到端拥塞控制方案都属于被动控制，即源端在拥塞发生前总是盲目地提高发送速率，而发送速率的提高又可能导致网络中排队的产生，形成拥塞。拥塞情况反馈到源端后，源端才会被动地降低发送速率。但是，由于反馈总是存在一定的时延，并且不能准确反映网络的拥塞程度，因此被动拥塞控制总是会导致网络在拥塞和非拥塞状态之间震荡，使网络带宽无法得到充分利用，同时较深的队列排队造成了较大的长尾时延，对时延敏感的I/O影响很大。端网协同主要根据网络可用带宽，准确调整端侧发送速率，这种源端与交换机之间的密切配合使网络中的队列近乎为空，同时能保持接近100%的带宽利用率。

在RDMA传输场景，根据实验室测试，采用典型拥塞场景哑铃状拓扑，与当前业界主

主流的协同算法相比，端网协同算法能够把排队时延控制接近于0。

<b>N=200</b>	<b>端网协同</b>	<b>HPCC</b>	<b>DCQCN</b>
<b>50%-ile</b>	0.155	3.023	116.612
<b>90%-ile</b>	0.238	6.662	121.82
<b>99%-ile</b>	0.321	8.204	125.48
<b>99.9%-ile</b>	0.401	9.094	127.131

表4 不同拥塞算法排队时延对比

在TCP传输场景，根据实验室测试，端网协同算法相对于传统的TCP cubic算法、bbr算法，可以做到0丢包，避免了现网因为TCP慢启动机制造成的流量突发丢包，保证了端侧APP的性能；同时，由于CAQM控制了队列的深度，有效降低了TCP所在队列的时延，避免了长尾时延对端侧业务性能影响。实验室实测数据表明，相对于传统TCP cubic拥塞控制算法，端网协同算法可以减少90%左右的长尾时延。

### 4.2.3.2 FlexBuffer拥塞控制技术

DCN的流量调度策略需充分考虑流大小和数量的差异性。在大量小流少数大流的场景中，采用“小流优先、大流后传”的调度机制可显著降低平均FCT。其核心原理在于：若大流优先传输，将阻塞其后到达的大量小流，导致整体等待时间增加；而若优先调度小流，虽可能略微延长少数大流的传输延迟，但由于小流数量远多于大流，整体平均FCT得以有效降低，系统效率更优。

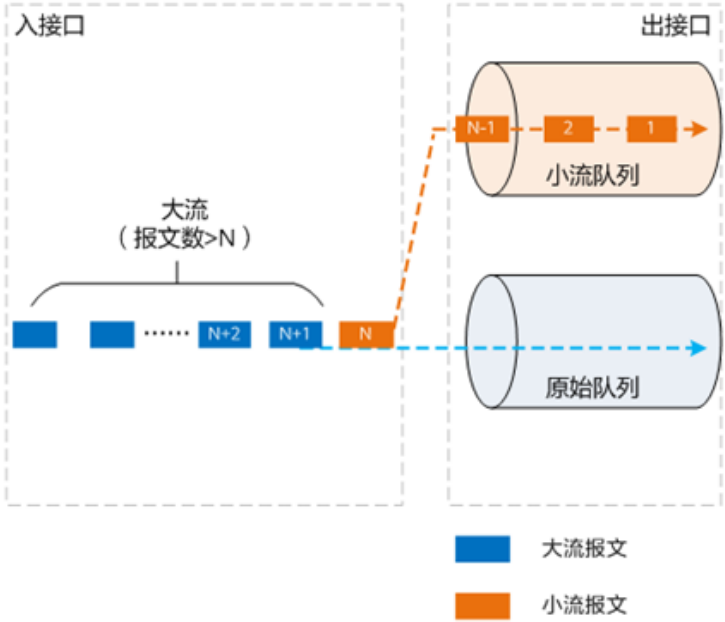


图18 大小流报文调度机制

为此，网络设备需实现基于流大小的差异化调度。在流建立初期判断其为大流或小流，并将其分配至相应的优先级队列：小流进入高优先级队列，大流则进入低优先级队列。调度器采用严格优先级调度策略，即高优先级队列中的所有流必须完全转发后，才开始处理低优先级队列中的流量。该机制显著缩短了小流在网络中的排队与传输延迟，避免其被大流“尾部阻塞”，从而有效提升整体业务响应速度。

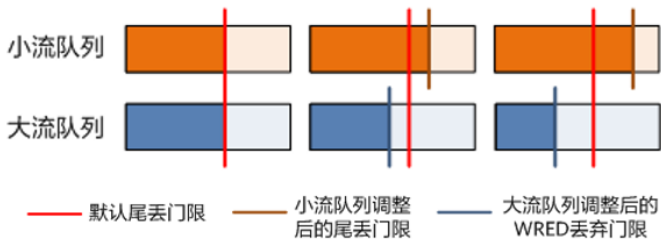


图19 基于流大小的差异化调度

在实现大小流差异化优先级调度的基础上，可进一步对高、低优先级队列采用精细化的缓存管理策略，以更好地匹配不同类型流量的传输特性与服务质量需求。小流通常具有突发性强、完成时间敏感的特点，其性能高度依赖端到端时延。为保障小流快速完成，应为其分配相对充足的缓存资源，避免因队列拥塞导致排队延迟增加或丢包重传。相比之下，大流持续时间长、数据量大，更关注整体吞吐稳定性而非单个报文时延。因此，大流对缓存资源的需求相对较低，可配置较小的缓冲区配额。结合拥塞控制机制，通过及时的拥塞信号反馈与速率调节，即可实现大流的平滑传输，避免网络资源过度占用，提升链路利用率与公平性。

#### 4.2.3.3 拉远训练精准流控技术

在万卡乃至十万卡级大模型训练场景中，受限於机房空间、电力与散热瓶颈，集群不得不跨多个机房甚至地理局点部署。同时，模型迭代加速，算力需求激增，单一数据中心难以承载，多DC协同成为支撑更大模型训练的必然选择。此外，租户算力本地化需求及部署不均导致资源碎片化，如何整合分散算力、实现跨集群高效协同，成为提升整体算力利用率的关键。

相比本地部署，拉远训练面临显著的网络挑战：一是时延增加，每10km带来20 - 30ms RTT增长，影响集合通信效率；二是长距链路易丢包，千分之一丢包即导致RDMA吞吐下降1/3。尤其在光模块闪断、链路误码或服务器拥塞等异常下，跨DC链路缓存不足，易引发严重丢包，使迭代时间翻倍。

具体来说，多DC训练的拥塞发生的位置不同，带来的结果也不同。如果拥塞发生在数据中心内部，则无论是流控还是拥塞控制，都可以很快生效，抑制住拥塞；而如果拥塞发生在跨长距链路上，因为反馈时间长，设备缓存不足以接纳链路在途数据包，从而发生丢包。网络交换机可以通过精准流控技术有效解决这个问题。

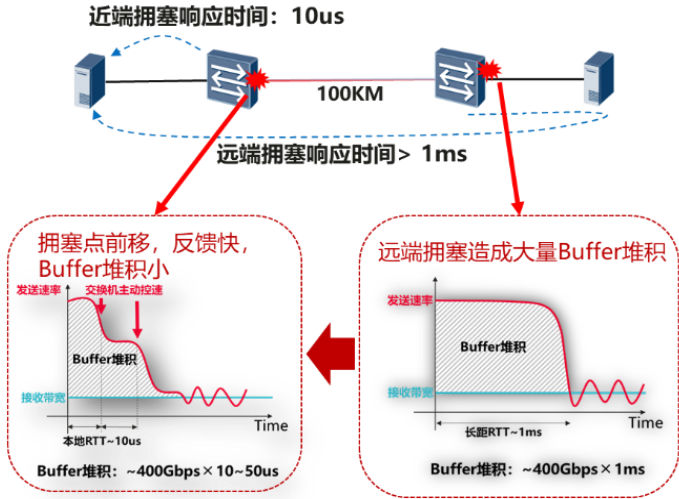


图20 拉远训练精准流控

在拥塞发生前，将长距链路上的潜在拥塞“前移”至源端Leaf交换机进行限速。通过全网状态监测，拥塞点向上游通告，源端依据算法动态调节流量，并通过PFC/CNP等机制实现精准控速。

结合智算流量周期性特征——每轮迭代通信模式相似，系统可在首个周期记录拥塞流，后续周期主动限速，实现“一次感知、周期无损”。该机制显著降低跨域丢包风险，保障拉远训练的稳定与高效。

#### 4.2.3.4 AI ECN 2.0技术

由于稀疏模型的兴起，网络拥塞成为影响模型训练的重要因素。引起拥塞的原因有：ALL TO ALL类集合通信的接收端多打一；MOE多种并行类型并发时对网络通信带宽的争抢；多任务混跑共交换设备上下行链路。

DCQCN是当前主流网络设备都支持的拥塞控制技术，华为在DCQCN基础上研发了AI ECN技术，动态调节ECN参数来实现动态无阻塞，被广泛应用在数据中心场景。

随着交换机代际演进，缓存增长速度落后于交换容量和端口速率的增长速度，导致传统ECN先触发降速，PFC兜底无丢包的拥塞机制失效。

因此，更加智能的的ECN参数调节方式成为解决AI网络拥塞的技术途径之一。

## 4.2.4 在网计算技术

传统数据中心网络被视为“透明管道”，仅负责数据的转发与路由，所有计算任务均由服务器端完成。而在大模型训练、分布式推理等高性能场景中，频繁的集合通信（如AllReduce、AllGather）会产生大量中间数据传输，造成网络拥塞和端侧处理瓶颈。在网计算通过在交换机或网卡中嵌入可编程处理单元，在数据流经网络设备时实时执行特定计算任务。例如，在分布式训练中，多个GPU需将梯度数据汇总并平均，传统方式是将所有梯度传至某节点集中计算，而在网计算可在交换机中逐跳或分层完成梯度聚合，仅将最终结果传回终端。在网计算技术不仅可以减少传输数据量，而且可减少通信次数，从而提高集合通信效率并加速应用性能。

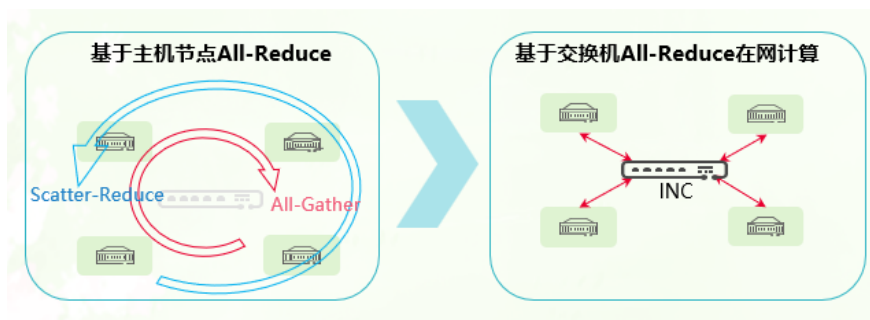


图21 在网计算技术

在网计算是典型的端网协同技术应用场景，该技术的使能需要端侧和网络侧的密切协同以及硬件与软件的紧密配合；在执行在网计算的过程中不仅需要网络设备的聚合卸载处理能力，也需要端侧软件栈感知和协同在网计算任务，以保证计算过程的正确性、完整性和可靠性。端侧节点和交换机节点通过一个统一的控制面管理器实现整个在网计算流程的管理和控制。

## 4.2.5 在网存储技术

在大模型推理场景中，海量且多样的用户请求导致Prefill和Decode阶段生成的KV Cache数据量巨大，超出DRAM存储容量，同时频繁的KV Cache传输引发RDMA网络流量竞争，影响推理性能。现有Prefill-Decode分离架构忽视了KV Cache跨实例传输带来的通信开销，难以在有限的存储与网络资源下保障效率。

为优化这一问题，可在交换机或专用服务器中部署智能KV Cache管理机制。通过监控报文交互和内存扩展情况，可识别系统是否利用外挂内存进行KV Cache调度。基于“以存代算”和“以查代算”技术，系统利用流计数器统计每个KV Cache块的访问频率，包括源实例、目的实例及传输内容，进而识别热点KV Cache并分级管理：

第一级为低频热点，被少于5%实例访问多次，仍存于本地资源池；

第二级为中频热点，被超过5%实例偶尔访问，由交换机或专用服务器缓存，按需下放；

第三级为高频热点，被广泛频繁访问，系统通过组播方式将其分发至所有未拥有的实例，实现全集群共享。

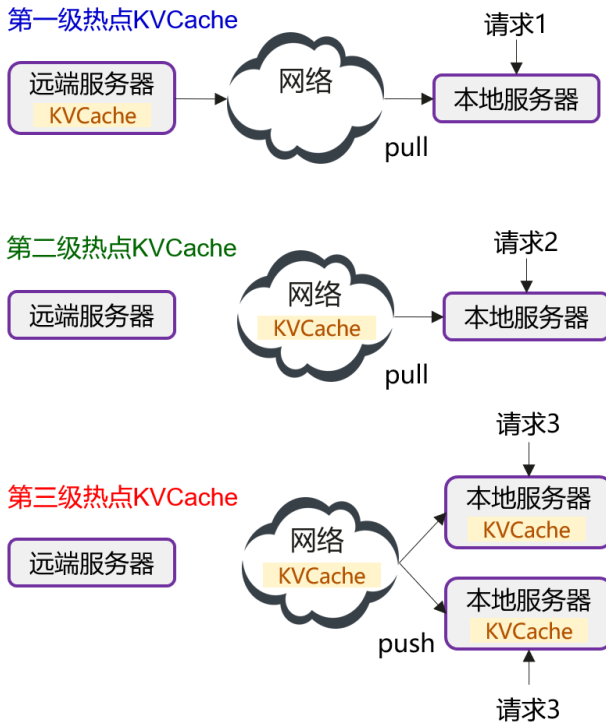


图22 在网存储技术

当KV Cache流经定制化交换机或服务器时，系统实时提取块信息，一旦判定达到第二或第三级热点标准，即截取并卸载至中间节点存储。该方法有效减少重复传输，缓解网络拥塞，提升存储利用率和推理吞吐。

## 4.2.6 高稳韧性技术

### 4.2.6.1 故障恢复技术

当链路故障发生时，传统路由收敛技术依赖控制面的动态路由协议（如OSPF、BGP）的信息交互和重新算路，即使结合了BFD探测技术也只是加快了故障感知速度，路由收敛时间仍在百毫秒级，大规模DCN网络中的路由收敛时间甚至达到秒级。对于在线交易类型的应用，比如高性能存储业务或高性能数据库访问业务，都需要极致的性能和高可靠性保障。针对此类业务，一旦链路故障发生，传统收敛技术在百毫秒内无法恢复正常业务传输，持续丢包会导致交易失败，甚至可能引发对端协议栈连接超时，应用性能会出现明显大幅下降。

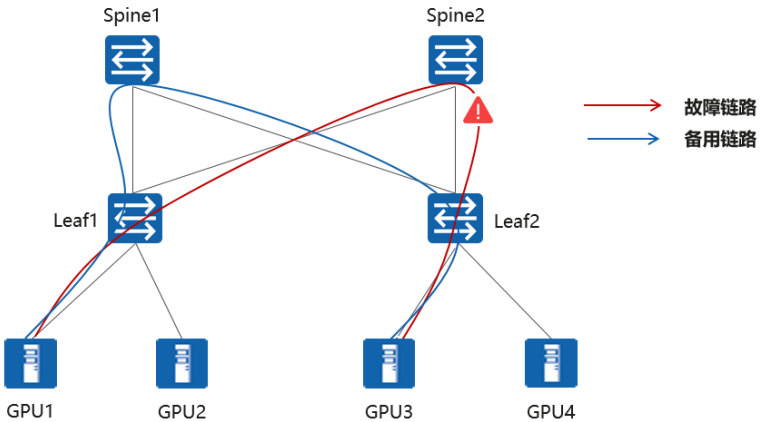


图23 组网中故障链路切换

在典型的Clos组网中，交换机之间都有多条路径，当一条链路出现故障的时候，通过感知端口状态、路由收敛、转发路径切换等操作，完成流量从故障链路到备用链路的收敛。但是这个时间一般在秒级。然而在AI场景里面，每次通信时间在毫秒级别，秒级时间内正常情况下已完成了多轮通信。如果依靠传统的路由收敛方式，将极大的影响AI计算效率。

#### 数据面快速恢复技术

数据面快速恢复技术在此场景下，可以做到毫秒级收敛，提供基于数据面的故障快速感知和快速收敛能力，提升设备级可靠性。特性包含故障快速感知，故障本地快速收敛，故障

通告生成、接收和中继处理，故障远程快速收敛和表项老化处理。针对在线交易类关键应用场景，尽量做到应用无感知的故障快速收敛效果，即在链路故障发生时业务性能无明显下降。

下面通过以下两个典型场景介绍数据面快速恢复技术的功能特性：

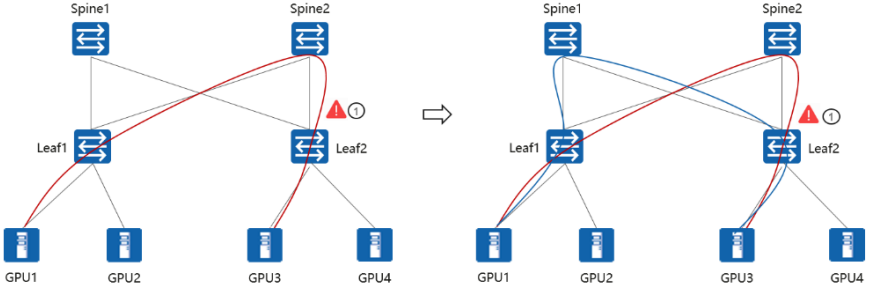


图24 场景1本地快速收敛

此场景的特征是有多个下一跳。比如上图中的GPU3访问GPU1的流量，到达leaf2以后，leaf2的下一跳可以选择spine1或者spine2。当leaf2->spine2的路径发生故障时，可以在本地快速切换到leaf2->spine1路径。

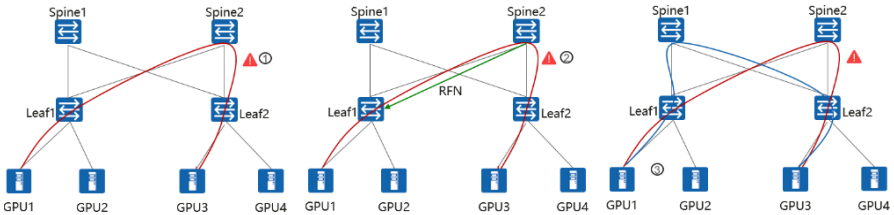


图25 场景2远端快速收敛

此场景的特征是有一个下一跳。比如上图中的GPU1要访问GPU3，路径是 leaf1->spine2->leaf2。当报文到达spine2以后，实际只有一条链路可以到leaf2。在此情况下，如果spine2->leaf2 路径出现了故障，spine2会通过发送通知报文给leaf1，leaf1在本地完成路径切换，切换到leaf1->spine1->leaf2路径。

## 数据面故障穿越技术

当网络中出现静默故障导致的业务会话级异常，如链路故障，转发表项异常，转发器件异常、端口物理状态是UP但无法转发流量、配置错误等，往往无法快速识别。当前业界主要通过分析器采集流量，检测异常，然后人工介入逐一排查，修复故障。排障需要数十分钟到数小时不等，严重影响业务。在这种情况下，华为进一步推出数据面故障穿越技术，提供基于数据面的故障快速感知和快速收敛能力，最快达成秒级路径切换，提升网络级可靠性。无需分析器和人工参与，自动感知网络故障，涵盖路由黑洞、ARP表项异常等业界无法自动检测的故障场景，然后根据预置策略换路，实现秒级路径切换。在金融、存储以及超算等关键场景，自动恢复故障,保障业务秒级恢复。

如下图所示，TCP流量传输过程中，发送端发送数据报文给接收端，接收端成功接收到数据报文后，会回复给发送端一个ACK报文，表示确认接收到ACK确认号前面的所有数据。如果超过一定时间未收到ACK报文，发送端设备会重传数据报文。设备针对ACL规则匹配的TCP流量建立流表，如果设备重传数据报文的时间间隔超过用户设定的故障检测时间，则判定超时故障，从而实现网络故障感知功能。

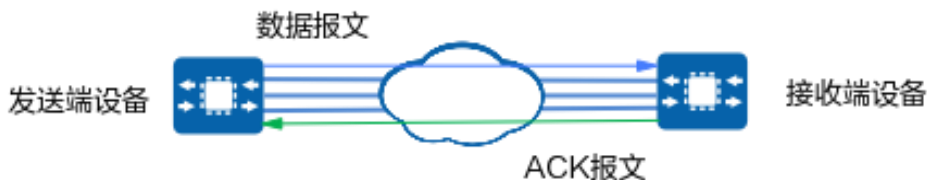


图26 TCP的确认机制

如下图所示的传统L3组网，服务器采用独立IP接入，Leaf交换机部署为独立的L3网关，承担二、三层流量转发，Spine层部署为独立的L3设备，与Leaf之间形成ECMP，实现流量负载分担。这种组网主要应用于金融、存储、超算等无损场景。网络中一旦出现静默故障，中断时间长，对上层业务影响严重。比如对于在线交易类型应用，如果出现持续丢包会导致交易失败，甚至可能引发对端协议栈连接超时，应用性能会出现明显大幅下降。部署该功能后，当业务流转异常时，设备能够自动感知故障，同时快速切换到其他ECMP成员，保障业务秒级恢复。

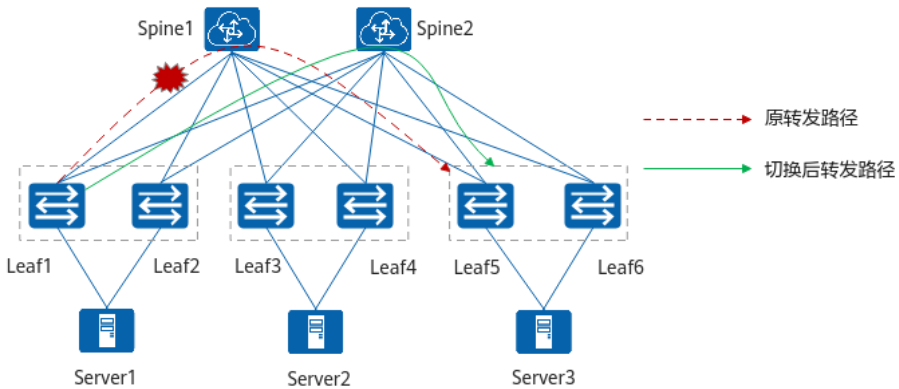


图27 传统L3组网

#### 4.2.6.2 闪启技术

万卡集群单设备部署，组网中leaf和NPU间无备份链路。同时集群断训代价大，无升级和维护窗口。业界leaf故障或升级导致的重启断链时间>120s。而当前昇腾NPU建链超时时间是30s。即，一旦设备异常重启，就会导致训练任务中断。接入交换机故障，影响所有直连NPU卡，需回退到checkpoint点，重新训练。中断一次重新拉齐，平均浪费2小时训练成果，仅电费成本一项¥10w+。

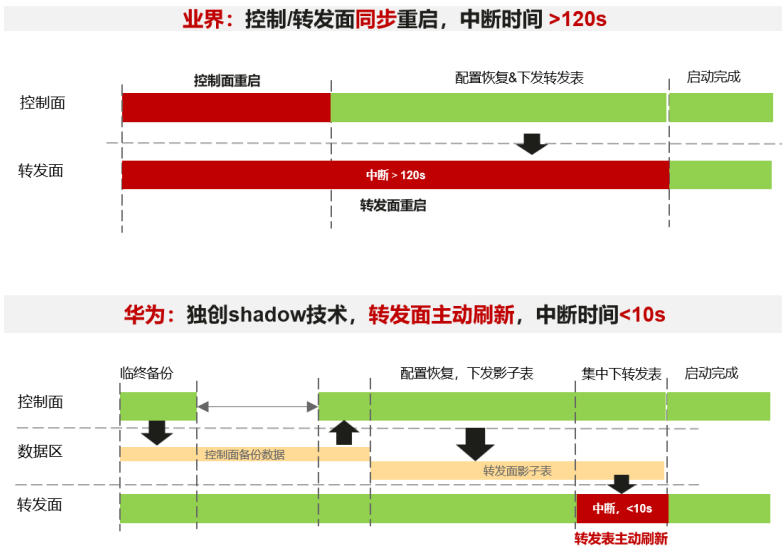


图28 闪启与业界实现对比

华为提供闪启功能，提升单点设备可靠性。在故障重启和升级重启场景下，实现流量中断时间<10s，保障训练任务不中断。重启期间，系统保持接口不Down，转发芯片不重启，流量正常转发。控制面重启后，DB等机制侧数据恢复，把数据从持久化存储设备中导入到内存。系统各组件启动、恢复，在全平滑过程中，控制面数据刷新到shadow表中。系统重启后，短暂中断硬件转发，将全量shadow数据搬迁至硬件转发层面，随后立即恢复业务。

通过闪启功能，可减少异常场景导致的业务中断，也提供了升级不断训的维护手段。可提升客户网络的单设备可靠性和可维护性。

### 4.2.6.3 光链路检测技术

#### 光模块通道抗损

在AI训练中，成千上万张算卡协同完成一项任务。一旦出现单点故障，整个训练任务都会被迫中断，而光模块是保证AI训练稳定性的关键一环。传统光模块的年失效率高达4‰，这导致万卡集群每年因光模块失效中断训练约60次。其中，单通道故障约占90%。频繁的训练中断不仅严重影响训练效率，还增加了维护和时间成本。对此，光模块通道抗损技术可显著降低光模块故障率，确保训练任务的连续性，从而保障系统的高效运行，提升网络的整体可靠性和稳定性。

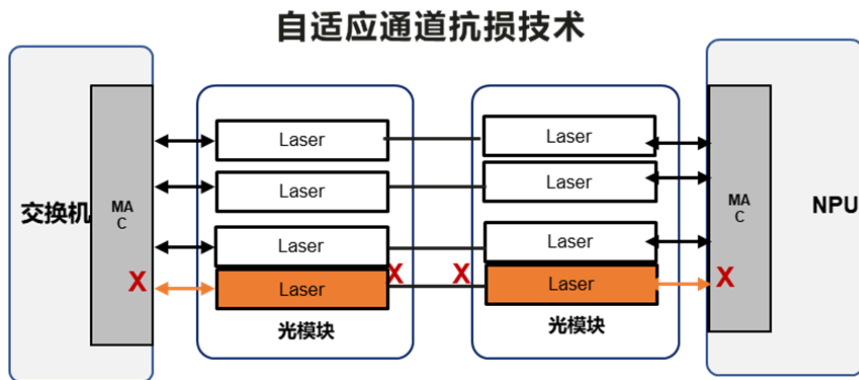


图29 自适应通道抗损技术

通过自适应通道抗损技术，可以有效解决传统光模块因单通道故障导致的训练中断问题。当单个通道故障时，通过交换机和NPU协商探测，精确识别出故障通道信息。仅对故障所在通道组停止工作，其他通道组仍可正常进行数据转发。通过单lane隔离，光链路降速不中断。避免故障链路切换增加其他路径带宽负荷，影响网络吞吐。

通过光模块通道抗损技术，光模块年失效率从4‰降低至0.4‰，万卡集群每年因光模块失效导致的训练中断由原来的60次减少至6次，网络稳定性提升10倍。



图30 华为光模块与业界对比

### 光模块脏污/松动检测

对于光模块故障问题，传统运维无法感知光链路脏污/松动，只能在光模块彻底故障、所承载的业务完全不可用时，才能感知到故障。属于故障事后感知，发现时已经造成业务中断。尤其在智算场景中，智算集群中断1小时，客户经济损失严重。同时由于光模块长期运行，光器件性能逐渐衰减，在光模块的生命周期中有一段时间处于亚健康状态。亚健康状态的链路不稳定，容易出现光模块闪断引起训练频繁中断。

为此，需要提供光模块的故障主动检测能力。基于采集到的光模块收/发光功率及门限阈值、偏置电流、电压、温度，并结合端口管理态、运行态、端口错包等指标，持续分析光链路指标数据，对比标准值，实时检测光链路亚健康异常。在出现光链路脏污/松动时，1分钟定位出故障点，先于业务受影响前定位到异常光模块。通过分析光模块指标异常，识别脏污问题及光模块风险隐患，便于工程师提前规划，替换掉亚健康状态的光模块，减少训练任务异常中断，提升连续训练时长。

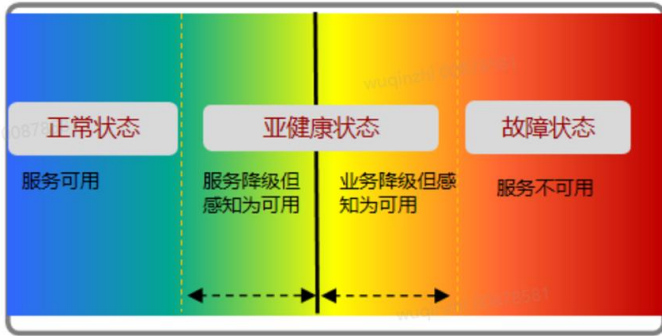


图31 光模块脏污/松动检测原理

### 光模块故障预测

结合大数据以及机器学习算法对光模块进行故障检测以及故障概率预测，先于业务受影响前识别出异常光模块。

基于采集到的光模块收/发光功率及门限阈值、偏置电流、电压、温度，并结合端口管理态、运行态、端口错包等指标进行数据训练，借助特定的故障模型以及线性回归、逻辑回归等AI算法识别已经发生的故障光模块，以及正在发生指标劣化但尚未发生故障的光模块。这样用户就可以在业务受影响之前，替换掉将会发生故障的光模块。

### 4.2.6.4 网络互联冗余

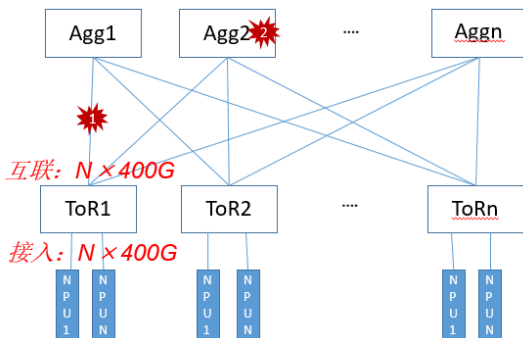


图32 互联侧故障会导致ToR上行带宽不足拥塞

随着AI集群向高密、大集群的方向演进，在同一个POD池内采用两层CLOS无收敛组网，该组网在接入和互联侧带宽是一致的，在hash不均匀，链路或者SPINE设备故障时，ToR的上行互联带宽<接入带宽，在ToR上行导致拥塞，从而影响整个集群的性能。

### 互联侧网络的降速风险:

随着交换机Radix的不断增加，集群规模越来越大，故障发生的概率也越来越高，互联侧降速运行的比例会增加。

序号	集群大小	单链路失效率	MTBF(H)	光链路修复时间(H)	降速运行时间比例
1	万卡	1000	98	4	4%
2	十万卡	1000	10	4	29%

表5 不同集群规模下，链路失效导致的降速运行比例

### 网络互联的冗余方案:

通过N+K的组网方式，在互联侧提供K条链路的带宽冗余，链路/设备故障点不多于K时，互联带宽≥接入带宽，可以保障互联故障后集群性能不降速。

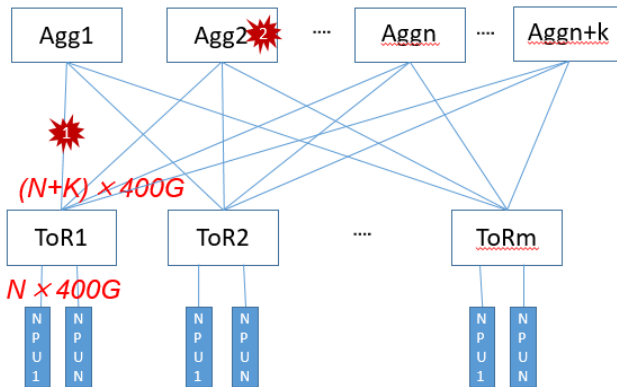


图33 N+K冗余组网故障后TOR上行不拥塞

## 4.3 AI 大脑

AI大脑依托AI模型训练与多智能体协同，提供智能化运维。网络智能体包括变更Agent、故障Agent、自动化Agent、网维Copilot等。

网络数字地图采集网络、流量、设备、服务器、应用等信息还原全网拓扑，可以实时呈现网络状态和业务路径，精准自动识别网络故障，实现隐患的深入分析与预防，让网络故障率降低90%。同时，通过业务仿真校验功能，预先评估实施方案的风险，确保网络配置变更100%正确。网络智能体大幅降低运维人员的技术门槛同时提升运营决策效率。

### 4.3.1 仿真验证/孪生仿真

**应用上线或变更，通常涉及大量网络变更，成为网络故障发生的主要根因。**以中大型金融数据中心为例，年均网络变更3000次以上，网络团队大量人力投入到变更设计和评估中。随着网络规模越来越大，网络配置越来越复杂，人工很难完全准确评估变更会给网络带来的影响，搭建实际的测试环境进行验证，也无法与现网生产环境1:1一致。

**风险排查与故障演练，通常涉及网络故障的覆盖，人工无法做到全覆盖。**随着网络规模越来越大，靠人工难以完成所有节点和链路故障的分析，动辄需要十几人投入一两周甚至一个月；另一方面，故障演练虽然定期执行，全部故障点遍历耗时久周期长，通常每次仅覆盖有限的几个故障点的演练，导致可能有潜在的风险没有被识别。

如何有效解决网络变更、风险排查、故障演练等场景评估难、耗时久的问题，网络仿真验证技术得到了学术界和工业界普遍的关注和实践。

网络仿真验证技术通常来说有两大类，Emulation方式和Simulation方式。

**Emulation方式：**基于虚拟化技术驱动设备镜像包实现，相当于在模拟器平台上搭建一套虚拟的网络环境，进行网络行为验证。其缺点是需消耗海量服务器资源：一台网元对应一个设备模拟器，数据中心一个中等规模分区200台网元，实现仿真推算需要20个VM。通常3节点的服务器集群，仅能支持几十台网络设备的组网模拟，仿真的网络规模无法满足当前大规模网络的要求。此外，仿真准确度还取决于模拟设备的镜像包与现网设备版本的一致性。需要设备厂家同步提供模拟器软件，且支持软转发，当前设备厂商支持度较低。

**Simulation方式：**即在采集现网配置、表项、拓扑等数据，并进行网络数据基线化建模的基础上，通过形式化验证或抽象图等算法，计算网络变更或出现故障点情况下，网络的

连通性与预期是否一致，并同时排查路由环路、路由黑洞、资源表项不足、协议两端配置不一致等潜在风险，最终给出仿真验证报告。

Simulation仿真技术由于其服务器资源消耗小、运行效率高，6000网元级规模，仅需1个VM。相比模拟器仿真分布式演算方案，集中演算消耗资源降低1000倍。可有效应用在网络变更效果评估、网络风险排查、网络故障演练场景，在大规模组网中，可有效提升网络变更评估效率和准确率，同时可将网络风险排查和网络故障演练，进行常态化仿真执行，加速排查周期，及时发现风险。

## 4.3.2 自动化Agent

在网络自动化领域，虽然各网络厂商的基础网络配置发放都已具备自动化能力，但随着网络规模变大、业务需求更加复杂，在网络自动化配置方面仍存在如下困难：

**多厂商网络设备共存**，网络变更通常会涉及多厂商设备，即便有统一的自动化工具平台，南向可以对接多厂商设备，但需要人工逐一编写适配不同厂商、不同款型的设备对接驱动包，编写驱动包工作量大，且对人员技能要求高。

在南向对接设备的驱动包准备好后，还需要**针对不同的业务变更诉求**，分别编写针对性的自动化工作流。同样，工作流的编写需要依赖人工，依赖编写人员对业务变更诉求、网络变更步骤、工作流编写技能，都有比较深入的理解。

**自动化Agent，内置大语言模型的学习理解能力，通过提前预训练和模型微调，使模型具有编写设备对接驱动包和编排工作流的基础知识能力。**在此能力基础上，通过对不同厂商不同款型设备的操作手册等资料进行学习理解，就可根据资料内容，自动生成用于设备对接的驱动包。类似的，通过对客户网络变更MOP文档、变更流程规范等已有的经验文档进行学习，结合已经生成的设备动作包，可将客户已有的变更流程生成自动化工作流，在后续遇到相同变更任务时，可直接执行工作流，提升网络变更效率。除此之外，面对新的业务变更诉求，自动化平台结合大语言模型能力，针对用户文字输入的变更意图描述，也可自动生成网络变更工作流，真正做到将运维人员的变更意图直接转化生成为网络变更工作流的能力。

基于大模型技术突破变更工作流的生成技术，提升作业自动化效率，降低Runbook流程设计的复杂度，提升业务变更开发的效率。

## 4.3.3 故障 Agent

当前，全球Top1000企业中的大多数仍受困于“告警风暴”：单日超百万条告警中仅10%左右具有实际处置价值，却需要消耗80%的运维人力资源。这种低效的运维状态背后，

折射出三重系统性矛盾：海量告警信息引发的“告警疲劳”导致关键事件漏判率高达30%；网络复杂度逐年增长趋势与人工处置效率逐年提升趋势间的剪刀差；知识库更新滞后造成的处置方案失效。传统“人工查阅文档-逐条执行步骤”的处置模式，使平均故障修复时间（MTTR）长期停留在30分钟左右，严重制约了业务的快速发展。

在数字化、智能化高速发展的背景下，网络运维正经历从“人工驾驶”向“自动驾驶”的范式转变。大模型技术与运维知识工程的深度融合带来了突破性变化。基于自然语言生成 workflow 模式，实现告警处置经验固化为自动化流程，从而实现故障处置效率的提升。

在数据中心网络运维领域大模型的发展形式正从自动化、智能化向多模态融合、平台化运维方向演进。尽管面临数据质量、可解释性和安全性等挑战，但其在数据&信息查询、故障检测等方面的应用已展现出巨大潜力。未来，随着技术的进一步成熟，大模型将成为数据中心运维的核心驱动力，推动行业向更高效、更智能的方向发展。

以金融行业为例，数据中心每日产生PB级的日志、流量和交易数据，性能指标远超人力分析极限，瞬时故障（如网络抖动、服务雪崩）需秒级定位，传统人工运维面临多系统割裂，信息孤岛现象严重，同时依赖运维人员自身的经验能力，往往需数小时排查，定位效率和分析能力与预期存在很大差距。相比之下，AI 大模型通过多模态数据融合分析，能够打破信息壁垒，将运维模式从被动响应升级为主动预测，有效满足金融业务对网络可靠性的高要求。

### 4.3.4 网维 Copilot

数据中心网络运维场景中，数据中心网络运维涉及告警平台、自动化平台、ITSM 等众多系统，各司其职却协同匮乏，形成一个个封闭的数据“烟囱”。这种跨系统协同机制的缺失，直接导致信息孤岛现象愈发突出，例如某条链路中断的告警触发后，运维人员需先在告警平台定位基础信息，再切换至自动化平台调取历史配置脚本，随后登录 ITSM 系统查询关联工单记录，全程需在 3-5 个系统间重复登录、筛选数据，仅全网数据检索环节就可能耗时 30 分钟以上。

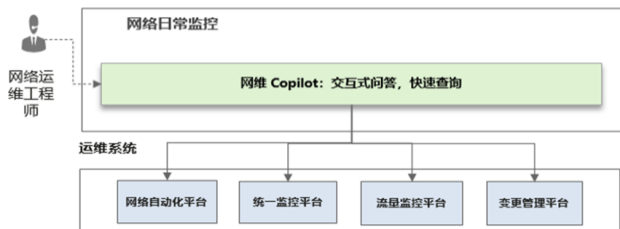


图34 大模型与运维系统对接

为解决上述运维痛点，如上图所示，利用大模型对运维工具的深度理解与智能调用能力，通过预先学习系统运维工具的 API 接口，构建工具能力知识库。当运维人员以自然语言发起查询时，大模型会精准解析意图，自动匹配最优 API 组合，完成跨系统数据的实时获取、清洗整合与多模态呈现。

# 5

## 总结和展望

随着数字经济向智能世界加速渗透，算力已成为核心生产力，而数据中心网络作为“算力流通血管”，其演进将进一步围绕“算网深度协同”展开，研究报告结合技术趋势与产业需求，提出三大展望方向：

### 1) 从“承载算力”到“调度算力”，网络成为算网协同核心枢纽

未来数据中心网络将突破“传输通道”的传统定位，依托 AI 大脑的全局感知能力，实现“算力供需匹配”：一方面，实时采集各节点算力负载、存储资源、业务优先级等信息，动态调度数据传输路径，避免“算力过载节点拥堵、闲置节点浪费”的问题；另一方面，将网络调度能力与 AI 平台深度融合，支持“算力按需分配”（如将高优先级 AI 推理任务调度至低时延节点），推动算力从“静态池化”向“动态流动”升级，最大化释放算力价值。

### 2) 技术持续迭代，向“更高速、更智能、更绿色”演进

在技术维度，AI Fabric 2.0 将持续突破性能边界与能效瓶颈：

**带宽升级：**伴随十万卡级 AI 集群落地，1.6T/3.2T 以太技术将快速商用，满足更大规模分布式计算的流量需求；

**智能深化：**AI 大脑将引入大模型技术，实现“故障根因自动定位、业务流量智能预测、网络配置自优化”，进一步降低人工运维依赖；

**绿色低碳：**通过芯片级能效优化（如低功耗转发芯片）、流量压缩技术、动态节能策略，推动数据中心网络 PUE（能源使用效率）进一步降低，契合“双碳”目标下的产业绿色转型需求。

### 3) 生态协同加速，构建“算网融合”产业共同体

AI Fabric 2.0 的价值落地需产业链各方协同：一方面，网络设备厂商将与服务器、AI 芯片厂商深度合作，推动硬件接口标准化与软件协议兼容性，避免“技术孤岛”；另一方面，将联合云服务商、AI 算法企业，针对自动驾驶、智能制造、生物医药等垂直场景，开发“算网一体”解决方案，形成“技术-场景-价值”的闭环。未来，“算网融合”将不再是单一技术或产品的创新，而是产业链共同参与的生态工程，推动 AI 产业从“单点突破”向“规模化落地”迈进。

综上，AI Fabric 2.0 不仅总结了数据中心网络的演进规律，更明确了AI时代网络的核心定位，网络架构重构涉及的关键技术路径；未来，随着算网协同的深化与生态的完善，数据中心网络将成为驱动智能产业发展的关键基础设施，为数字经济向更高质量迈进提供坚实支撑。



扫码了解更多