2025
AI网络技术白皮书

2025年8月

前言

人工智能(AI, Artificial Intelligence)技术,尤其是大模型(LLM, Large Language Model)技术的广泛应用,正驱动着 AI 与网络深度融合创新,成为产业变革的关键力量。据 IDC 报告,2024 年下半年,中国智算基础设施服务市场规模达 124.1 亿元,2028 年中国智能算力规模预计将达到 2781.9 EFLOPS,市场正从百亿爆发期快速迈向千亿成长期。AI 的广泛应用使得网络基础设施面临着超高带宽、超低延迟、确定性传输、端网协同及网络智能化等新需求,网络在架构、协议、运维和成本等方面遭遇严峻挑战。

面对这些新需求与挑战,本白皮书将 AI 网络划分为“网络赋能 AI(Network for AI)”和“AI 赋能网络(AI for Network)”两大方向。在 AI 网络的发展框架中,“网络赋能 AI(Network for AI)”以满足 AI 技术与应用的网络新需求为导向,通过纵向扩展(Scale Up)、横向扩展(Scale Out)以及确定性网络(DetNet, Deterministic Network)等前沿技术突破,为 AI 训练集群的高效运转、推理服务的稳定输出以及分布式 AI 应用的广泛落地提供坚实保障,让网络真正成为 AI 创新发展的坚实后盾。“AI 赋能网络(AI for Network)”则聚焦借助 AI 技术提升网络自身的智能化水平,其核心目标是实现网络的智能自治,通过意图驱动网络(IDN, Intent-Driven Network)、数字孪生网络(DTN, Digital Twin Network)、智能网络大模型等技术的深度应用,简化网络管理操作流程、提升网络运行效率、强化网络防御能力,推动网络从传统运维向自主智能方向加速演进。

“Network for AI”和“AI for Network”双向赋能和协同创新,不仅是突破网络技术瓶颈的关键,也是推动 AI 规模化应用的引擎,驱动着网络、算力、数据等产业要素融合升级,更将重塑金融、制造、医疗等垂直行业的数字基础设施形态,推动智能社会建设进入快车道。

在此背景下，江苏省未来网络创新研究院联合硬件厂商、软件厂商、服务商和行业应用企业等产业链伙伴共同编写了本白皮书，旨在通过系统梳理 AI 网络技术的演进、现状与展望，凝聚产业共识，助力网络与 AI 技术深度融合、产业生态健康发展，为智能时代的网络基础设施建设提供系统性指导与前瞻性思考。

编写委员会

专家指导单位：北京邮电大学、紫金山实验室

主编单位：江苏省未来网络创新研究院

参编单位：中国铁塔股份有限公司、天翼云科技有限公司、奇异摩尔（上海）半导体技术有限公司、无锡沐创集成电路设计有限公司、益思芯科技（上海）有限公司、深圳第一线通信有限公司、江苏致网科技有限公司、苏州衡天信息科技有限公司

媒体发布：究模智、SDNLAB

目录

第三部分 未来展望

第 1 章　AI 与网络的融合演进

AI 与网络的融合经历了初步探索、快速发展、深度融合的演进过程，每个阶段都有其独特的技术特征和产业标志，构成了 AI 网络技术演进的完整脉络。

第一阶段：初步探索期

AI 与网络的融合始于深度学习（DL，Deep Learning）的突破性进展。2012 年，深度学习先驱、AlexNet 之父 Alex Krizhevsky 及其团队使用 2 块 GTX 580，耗时 5-6 天训练的 AlexNet 模型在 ImageNet 竞赛中夺得冠军，这一事件标志着深度学习在图像分类领域的重大突破，开启了深度学习快速发展的新篇章，也首次显性暴露出单机算力瓶颈，激发了对分布式训练网络（DTN，Distributed Training Network）的迫切需求。2012 年左右，谷歌 B4 网络投入使用，推动了网络可编程化的实践。2016 年，AT&T 发布 Domain 2.0 计划，进一步加速了软件定义网络的规模化落地，为后续 AI 驱动的智能网络埋下伏笔。该阶段以 AI 算法创新为核心，网络主要承担着基础数据传输功能。主要技术特征表现为：

- AI 侧：聚焦 CV/NLP（Computer Vision / Natural Language Processing）领域的模型革新（如 AlexNet、ResNet），模型规模较小，训练依赖单机或小规模 GPU（Graphics Processing Unit）集群；
- 网络侧：仍以传统 TCP/IP（Transmission Control Protocol/Internet Protocol）协议为主，数据中心采用胖树架构（Fat-Tree），带宽普遍≤10Gbps，远程直接内存访问（RDMA，Remote Direct Memory Access）技术初步尝试但尚未普及。

关键挑战：网络性能严重制约 AI 发展，分布式训练中参数同步耗时占比极高，时延问题和带宽不足导致训练效率低下；AI 技术尚未应用于网络优化，运维完全依赖人工。

第二阶段：快速发展期

大模型时代的到来驱动 AI 与网络进入协同升级的快车道。2017 年由 Google 团队提出的 Transformer 是指一种基于“自注意力机制”（Self-Attention）的神经网络架构，如今在 AI 领域得到广泛应用，随后 GPT、BERT 等模型参数量突破亿级，千卡协作训练需求倒逼网络低时延与高吞吐能力跃升。2020 年，英伟达收购 Mellanox，整合 GPU 算力与 InfiniBand 网络技术，端到端 AI 架构奠定了超算网络的新标准，也迈入了 AI 与网络快速融合发展的新阶段。该阶段的技术特征呈现协同发展趋势，主要表现为：

- 网络赋能 AI 发展：数据中心向高扩展性多级交换网络架构演进，100G/200G 带宽普及，RDMA 规模化部署将时延压缩至微秒级；确定性网络和无损算法保障 AI 集群通信质量。

- AI 赋能网络智能化：AI 技术开始应用于智能流量调度、网络故障预测与定位、资源分配等场景，初步实现部分运维自动化，提升网络管理效率和韧性。

关键挑战：超大规模场景成为关键制约因素。网络性能难以适配 AI 规模扩张需求，万卡集群中通信效率大幅下降，影响训练进程推进；AI 对网络的智能化赋能范围有限，面对复杂动态的网络环境，仍难实现全面高效的自动化管理与优化。

第三阶段：深度融合期

AI 与网络技术已迈入双向赋能的闭环发展阶段，“AI 网络”正逐步确立为核心技术范式。2022 年，ChatGPT 的全球爆发成为重要转折点，千亿参数规模的大模型对超低时延响应速度的刚性需求，直接驱动了低时延推理网络的加速建设。在此背景下，国内外云服务企业及大型科技公司纷纷加大投入，布局专用智算中心。国家层面亦高度重视，通过系统性规划推进 AI 智算中心建设，以此支

撑大模型训练与推理过程中产生的海量算力需求，为人工智能技术的持续突破筑牢基础设施根基。这一阶段的 AI 网络呈现如下特征：

- 服务于 AI 的网络技术：新型网络架构和互连技术不断涌现，以满足极致性能需求。例如，英伟达的 InfiniBand 网络凭借超低时延和高吞吐性能，成为大规模 AI 训练集群的首选；以超以太网联盟（UEC，Ultra Ethernet Consortium）为代表的开放组织也在积极推动以太网技术革新，致力于在以太网基础上实现媲美 IB 的无损、低时延特性，为行业提供更灵活的技术路径选择。

- 网络智能化技术：意图驱动网络逐步走向成熟，数字孪生网络可实现网络状态的精准感知与全生命周期管理，智能运维应用开始渗透网络全场景，实现网络配置、优化、故障处理的闭环自治。

关键挑战：800G/1.6T 的网络互连在物理性能和成本上难以满足万亿参数模型的训练需求；数字孪生网络因数据不互通，影响了对网络状态的准确判断；网络自主管理能力跟不上 AI 流量的突然增长；智能运维的决策过程不透明，导致出现重大故障时还得依靠人工解决。

AI 与网络的协同发展呈现出“需求牵引、技术反哺、生态融合”的螺旋上升态势。首先，需求牵引成为技术突破的原始动力。随着 AI 模型参数从百万级飙升至万亿规模，对网络性能提出了极致要求，推动网络带宽在十年间实现近百倍增长，时延从毫秒级压缩至微秒级，甚至在工业控制等确定性场景下达成亚微秒抖动，直接驱动网络性能实现跨越式升级。其次，技术反哺形成产业发展的正向循环。在需求驱动下诞生的意图驱动网络、智能运维等网络智能化技术，反过来为网络提供强大支撑。不仅显著提升了网络资源利用率与智能化水平，更实现了运维成本的大幅降低，形成技术与产业相互促进的良性循环。最后，生态融合构筑智能时代的底层基石。跨域生态的深度融合与标准体系的持续完善，正为 AI 网络产业搭建起坚实的技术基础，为整个生态的可持续发展提供体系保障。

第一部分 Network for AI：面向 AI 的新型网络基础设施

第 2 章　AI 驱动下的网络架构变革与性能演进

本章深入探讨 AI 工作负载对网络基础设施提出的新型需求，分析超大规模训练、高性能推理以及边缘场景下的网络架构演进方向，为后续技术探讨奠定基础。

2.1　AI 驱动下的新型场景

随着 AI 技术向万亿参数时代迈进，其应用场景正从集中式训练向分布式推理与边缘协同加速演进。这一进程不仅依赖算法与硬件的突破，更需网络架构的深度适配。

（一）训练场景

AI 模型训练通常涉及对海量数据集进行迭代式学习，以优化模型参数。大型语言模型等超大规模 AI 训练需要数千甚至数万个 GPU 协同工作，训练网络中存在巨大的“东西向”流量，即 GPU 之间频繁的数据交换和模型参数同步。任何微小的网络瓶颈或数据包丢失都可能导致训练时间显著延长，甚至影响模型精度，为了确保训练效率和模型收敛速度，网络必须提供极致的性能。

（二）推理场景

AI 推理是将训练完成的模型应用于实际场景，以生成预测或执行决策的过程。与训练不同，推理更侧重于快速响应和高效处理“南北向”流量，即用户请求与 AI 服务之间的交互。推理场景对网络的需求因应用类型而异，但普遍要求

低延迟以提供流畅的用户体验。推理工作负载可能部署在数据中心、云端或靠近用户的边缘设备上，需要网络能够灵活、高效地连接各种终端用户和 AI 服务。

（三）边缘场景

边缘 AI 是将 AI 能力下沉到数据源头附近，例如智能摄像头、工业传感器等设备上。这种部署模式旨在减少数据传输到中心云的延迟和带宽消耗，同时提高数据隐私和安全性。边缘 AI 的兴起使得网络不再仅仅是数据传输的通道，更是 AI 计算的延伸。边缘网络面临的挑战包括资源受限、连接多样化（如 Wi-Fi、5G/6G、LoRaWAN 等）以及复杂的部署环境。

2.2 AI 网络需求

在 AI 持续向更智能、更泛在的方向发展的过程中，网络已然成为串联各种应用场景的核心纽带。只有不断突破网络架构的技术壁垒，实现网络与 AI 算法、算力基础设施等的深度融合，才能充分释放 AI 在各领域的潜力。以下是超大规模 AI 网络的核心需求：

（一）高带宽与低延迟

AI 网络中，计算节点需频繁交换数据，对网络带宽需求呈指数级增长。同时，通信延迟也会直接影响计算资源利用率：

- 超高带宽：AI 模型训练涉及海量数据的处理和传输，如百万亿参数级别的深度学习模型，其训练过程中产生的数据量极为庞大。超高带宽的网络能够确保这些数据在计算节点之间高效、快速地传输，避免因带宽不足导致的传输瓶颈，从而提升整体训练效率。

- 超低延迟：AI 应用，特别是实时推理任务，对网络的响应速度有极高要求。超低延迟的网络能够确保推理指令在极短时间内得到执行，提升用户体验和

系统实时性。此外，在 AI 训练过程中，超低延迟也有助于减少通信开销，提高计算资源的利用率。

（二）高可靠性与稳定性

AI 网络的高可靠性与稳定性是确保业务连续性和计算效率的关键因素，网络故障或性能波动可能对训练任务、实时推理等造成严重影响，因此需要：

- 无损传输：在 AI 训练过程中，数据包的丢失会导致模型训练中断或重新计算，严重影响训练效率和结果准确性。AI 网络需要实现无损传输，通过智能拥塞控制和流量调度机制，减少丢包和重传，保障数据完整性和一致性。

- 可靠性与稳定性：AI 网络应具备动态负载均衡能力，根据实时流量和网络状态，智能分配带宽资源，避免网络拥塞和性能下降。对于对网络实时性要求极高的 AI 应用而言，AI 网络需要提供亚毫秒级的故障检测和恢复能力，确保业务连续性。

（三）网络拓扑与通信模式

AI 模型和数据规模持续爆炸性增长，驱动计算集群从数千卡向数万、数十万卡扩展，网络拓扑必须具备灵活的扩展能力，并且与训练框架、并行策略深度协同，优化通信效率：

- 灵活可扩展的拓扑：网络需要支持灵活、高效的拓扑结构（如 Fat-Tree、Dragonfly 等），并能根据不同的通信模式动态优化路径，最大化通信效率，减少拥塞。

- 适配 AI 通信模式：针对 AI 训练中常见的 AllReduce、AllGather、ReduceScatter 等集合通信模式，在硬件与协议层面进行专项优化。

（四）数据安全与 QoS

超大规模 AI 集群通常同时运行预训练、微调、推理等多个任务，需通过资源隔离与服务质量（QoS，Quality of Service）机制避免相互干扰：

- 资源隔离：在大模型训推过程中，数据可能会包含商业机密、用户隐私等敏感信息。网络必须提供强大的安全保障，包括数据传输加密、严格的访问控制、网络切片/虚拟化隔离（确保不同租户或任务互不干扰）等。

- 流量优先级：精准识别并优先处理同步报文、控制信令等关键流量，应用优先级队列、带宽预留等 QoS 策略和动态流量工程，防止拥塞，保障可预测的延迟和低抖动。同时需兼顾东西向与南北向流量的高效协同。

总体而言，AI 网络的需求是高带宽、低延迟、高可靠、可扩展的综合体现，需从拓扑设计、协议优化、硬件加速、容错机制等多维度协同，最终目标是让网络实现高效互联，确保计算资源的利用率最大化，同时支撑模型规模和集群规模的持续突破。

2.3 AI 集群网络拓扑

AI 集群网络是支撑大规模人工智能训练和推理任务的核心基础设施，其拓扑设计直接关系到 GPU 等计算单元之间的通信效率，进而影响整体性能。

2.3.1 Fat-Tree 拓扑结构

Fat-Tree 网络架构由于其高效的路由设计、良好的可扩展性及方便管理等优势，成为大模型训练常用网络架构。对于中小型规模的 GPU 集群网络，通常采用 Spine-Leaf 两层架构。对于较大规模的 GPU 集群则使用三层胖树（Core-Spine-Leaf）进行扩展组网，不过由于网络的层次增加，其转发跳数与时延也会相应增加。

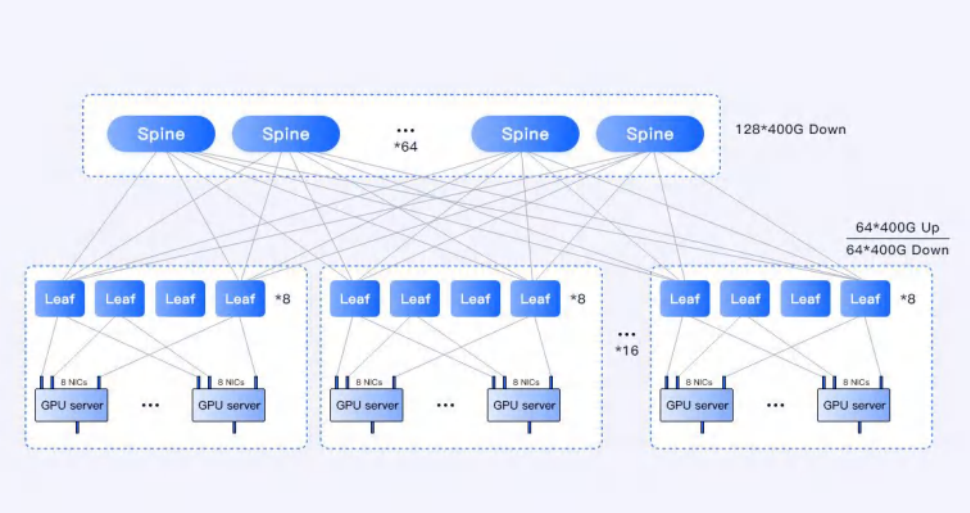

图 1 二层 Fat-Tree 架构

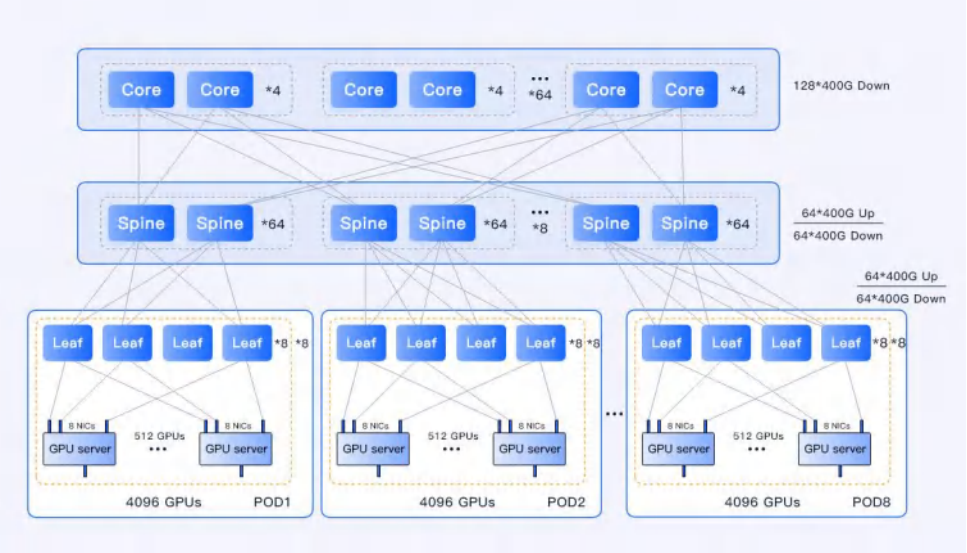

图 2 三层 Fat-Tree 架构

GPU 服务器的接入方式分为单轨与多轨两种。单轨接入方式是一台 GPU 服务器上的网卡全部汇聚于同一台 Leaf 交换机，集群通信效率偏低。多轨接入方式是将 GPU 服务器上的 N 张网卡各自接入 N 台 Leaf 交换机，集群通信效率较高。不过，若 Leaf 交换机出现故障，多轨方式下受影响的 GPU 服务器数量会比单轨方式更多。

Fat-Tree 架构的关键点在于交换机的上下行带宽的设计，保持 1:1 的无收敛配置，确保了网络中的任何流量路径都不会成为瓶颈，从而实现全带宽无阻塞通信。

Fat-Tree 的优势：

- 无阻塞通信：Fat-Tree 通过在网络高层提供足够的带宽，消除了传统树形拓扑中的带宽收敛问题，使得任意两个节点之间都可以实现线速通信，避免了网络拥塞。
- 高可扩展性：Fat-Tree 可以通过增加交换机和链路的数量来轻松扩展网络规模，支持数千甚至上万个计算节点的互联，满足大规模 AI 集群的扩展需求。
- 多路径冗余：任意两点之间存在多条等价路径，当某条链路或设备发生故障时，流量可以快速切换到其他可用路径，提高了网络的可靠性和容错性。
- 高吞吐量和低延迟：由于无阻塞特性和多路径选择，Fat-Tree 能够提供极高的数据吞吐量和极低的通信延迟，这对于 AI 训练中频繁的大规模数据交换至关重要。

Fat-Tree 架构作为当前 AI 集群网络领域应用最为普遍的主流拓扑结构，凭借其卓越的可扩展性与高效性，已成为众多大型项目优化部署与网络设计的基础架构。例如，字节 MegaScale 集群采用三层类 CLOS 拓扑，成功连接万卡级 GPU；阿里 HPN7.0 通过两层双平面架构，实现万卡级无拥塞互连。然而，Fat-Tree 架构也存在一定缺点，它需要大量的交换机和链路，成本较高，特别是对于超大规模集群。

2.3.2 Dragonfly 拓扑结构

Dragonfly 拓扑是一种为高性能计算优化的网络架构，通过减少网络直径来降低通信延迟。与 Fat-Tree 的对称设计不同，Dragonfly 采用分组全连接理念，更适应跨机柜、跨数据中心的 AI 训练场景。

Dragonfly 拓扑由多个组（Group）组成，组间和组内均建立全连接关系。每组间使用 1 条或多条链路连接，组内的每个网络节点都与组内其他网络节点直接互连，且可以同时连接到其他组和计算节点。

图 3 Dragonfly 架构

Dragonfly 拓扑中每个网络节点与同组内其他网络节点之间的带宽总和用 a 表示，每个网络节点与计算节点之间的带宽总和用 p 表示，每个网络节点与其他组之间的带宽总和用 h 表示。为了更好地实现负载均衡，建议满足 a=2p=2h。若使用其他值，则需满足 a≥2h 且 2p≥2h。

Dragonfly 的优势：

- 高可扩展性：Dragonfly 通过分层和全局互连的设计，能够以较低的成本扩展到非常大的规模，支持构建拥有数万甚至数十万个计算节点的超大规模 AI 集群。

- 低成本优势：相较于 Fat-Tree 等需要大量高端口密度交换机的拓扑，Dragonfly 通过减少全局链路的数量，显著降低了布线复杂度和网络设备的成本。
- 网络直径小：尽管是分层结构，但 Dragonfly 的平均跳数相对较低，有助于保持较低的通信延迟。

谷歌数据中心分布式交换架构 Aquila 采用了 Dragonfly 拓扑，其中核心创新在于：（1）全局链路优化：每个 Switch 通过少量高速链路（通常为光连接）与其他组的 Switch 直连；（2）最小化直径：确保任意两节点间最多 3 跳，组内 1 跳+全局 1 跳+目标组内 1 跳；(3) 虚拟通道：支持 8 个虚拟通道单独流控，避免队头阻塞。然而，Dragonfly 在 AI 计算网络中的应用相对有限，主要因为其软件成熟度不足且运维复杂度高，在流量突发情况下可能面临拥塞挑战。

第 3 章　面向 AI 的高性能网络关键技术

本章系统阐述支撑 AI 创新与规模化应用的高性能网络关键技术，全面剖析各项技术的原理、最新进展及创新实践，为构建新一代智算融合网络提供技术指南。

3.1　技术演进趋势与分类体系

3.1.1　技术演进趋势

AI 高性能网络的演进趋势主要体现在以下几个方面：

(1) 从通用网络向 AI 网络演进：传统网络主要面向通用计算和存储流量设计，其架构和协议难以完全满足 AI 训练和推理的极端性能需求。AI 工作负载的特点如大规模集合通信、高突发性流量和对时延抖动的敏感性，促使网络向 AI 专用化方向发展，包括采用更适合 AI 流量特性的网络拓扑、优化传输协议、以及引入更精细的拥塞控制机制。AI 网络将成为智算中心的核心基础设施，提供端到端的性能保障。

(2) 从硬件定义网络向软件定义网络与可编程网络演进：随着 AI 应用场景的日益复杂和动态变化，传统网络难以快速响应业务需求。SDN 和可编程网络技术为 AI 高性能网络提供了更大的灵活性和可控性。通过集中式控制器对网络资源进行统一管理和调度，可以实现网络路径的动态优化、流量的智能调度、以及网络资源的弹性伸缩。

(3) 从尽力而为向无损确定性保障演进：AI 训推对网络的服务质量有严格要求，任何性能波动都可能导致任务失败或效率低下。AI 高性能网络正在从传统的尽力而为向无损确定性演进。这包括对 AI 流量进行优先级划分、带宽预留、以及精细化的拥塞管理。例如，通过优先级流控制（PFC，Priority Flow Control）

和显式拥塞通知（ECN，Explicit Congestion Notification）等机制，确保 AI 训练流量在网络拥塞时仍能获得优先传输，降低丢包率和时延抖动，实现无损的数据传输，为 AI 任务提供稳定可靠的网络环境。

（4）从单一互联技术向多技术融合演进：为了满足 AI 工作负载对带宽和时延的极致要求，AI 高性能网络不再局限于单一的互联技术，而是趋向于多种技术的融合，Scale Up（如 NVLink、UALink）、Scale Out（如 InfiniBand、RoCEv2）技术正在从独立走向协同应用。不同技术在成本、性能和生态方面各有优势，通过融合应用可以构建出更具性价比和灵活性的 AI 网络解决方案。

3.1.2 分类体系概览

AI 高性能网络的技术可以从不同维度进行分类，以便更好地理解其构成和功能。本章将主要从以下几个核心维度进行分类和阐述：

Scale Up 技术：这类技术主要关注单个计算节点内部或紧密耦合节点间的性能提升，通过优化节点内/邻近节点间的数据传输效率，实现计算密度的指数级增长。其核心目标在于突破单机算力瓶颈，为集群提供超高速、低延迟的内部通信能力。典型的技术包括 NVLink、UALink 协议等。

Scale Out 技术：这类技术主要关注大规模计算节点的网络互连，通过优化集群拓扑、路由算法和传输协议，构建支持数万节点协同训练的分布式计算平台。其核心挑战在于平衡带宽、延迟、可靠性与成本，典型的技术包括 InfiniBand、RoCEv2 等。

前沿突破技术：这类技术代表了 AI 高性能网络领域的最新探索方向和突破性进展，不局限于 Scale Up 或 Scale Out 的框架，而是着眼于未来网络架构的革新、跨领域技术的融合以及全新通信范式的构建，为 AI 的未来发展提供强大的网络基础设施支撑。

3.2 Scale Up（纵向扩展）技术

3.2.1 高速互连技术

3.2.1.1 NVLink

NVLink 是英伟达开发的专有高速互连技术，旨在解决传统 PCIe（Peripheral Component Interconnect Express）总线在多 GPU 系统中的带宽瓶颈和延迟问题，从而实现 GPU 之间以及 GPU 与 CPU 之间的高效数据传输和协同工作。

（一）NVLink 的起源与演进

NVLink 最初于 2016 年与英伟达 Pascal 架构的 P100 GPU 一同发布，旨在为 GPU 提供比 PCIe 更高的带宽和更低的延迟。此后，伴随着英伟达 GPU 架构的每一次迭代，NVLink 也在不断演进，每一代都带来了带宽和性能的显著提升。

表 1 NVLink 的演进

	NVLink 1.0	NVLink 2.0	NVLink 3.0	NVLink 4.0	NVLink 5.0
推出时间	2016 (Pascal)	2017 (Volta)	2020 (Ampere)	2022 (Hopper)	2024 (Blackwell)
信号调制	NRZ	NRZ	NRZ	PAM4	PAM4
单链路带宽（双向）	20 GB/s	25 GB/s	50 GB/s	100 GB/s	200 GB/s
每芯片最大链数	4	6	12	18	18
拓扑扩展	点对点 (P2P)	NVSwitch 1.0	NVSwitch 2.0	NVSwitch 3.0	NVSwitch 5.0
协议改进	基础互联	支持 CPU-GPU 直连	链路级 CRC 校验	自适应路由 + 拥塞控制	亚微秒级同步协议
应用场景	HPC 初步集成	AI 训练加速	大规模 AI 集群	千卡级 AI 集群	万卡级 AI 集群

（二）NVLink 的技术特点

NVLink 采用多条高速差分信号通道组成链路的方式进行点对点通信。每个 NVLink 链路都提供双向数据传输能力，并具备极高的带宽。具体来看，从 P100

的 160GB/s 迭代至 B200 的 1.8TB/s，单卡带宽年复合增长率超 60%。NVLink 支持 GPU 之间直接进行内存访问，即一个 GPU 可以直接读写另一个 GPU 的显存，无需经过 CPU 作为中介，极大提高了数据传输效率，降低了通信延迟。此外，NVLink 支持多通道通信，允许同时进行多个数据传输会话。NVLink 不仅可以连接单个服务器内的多个 GPU，还可以通过 NVLink 交换机（如 NVSwitch）连接更多 GPU，构建更大规模的 GPU 集群，实现跨服务器的 GPU 互联，为超大规模 AI 训练提供强大的扩展能力。

（三）NVSwitch 全互联架构

NVSwitch 是在 NVLink 基础上发展起来的，它作为一个高速交换机，连接多个 NVLink，可在单一机架与多机架间以 NVLink 全速提供 GPU 完全通信。NVSwitch 支持完全无阻塞的全互联 GPU 系统，通过提供更多的 NVLink 接口，实现更大规模的 GPU 互联，从而构建更加强大的计算集群。例如，英伟达 NVLink5 Switch 具有 144 个 NVLink 端口，无阻塞交换容量为 14.4TB/s，能够支持多达 576 个完全互联的 GPU。NVSwitch 的出现，使得 AI 和 HPC 工作负载能够更高效地利用多 GPU 的并行计算能力。

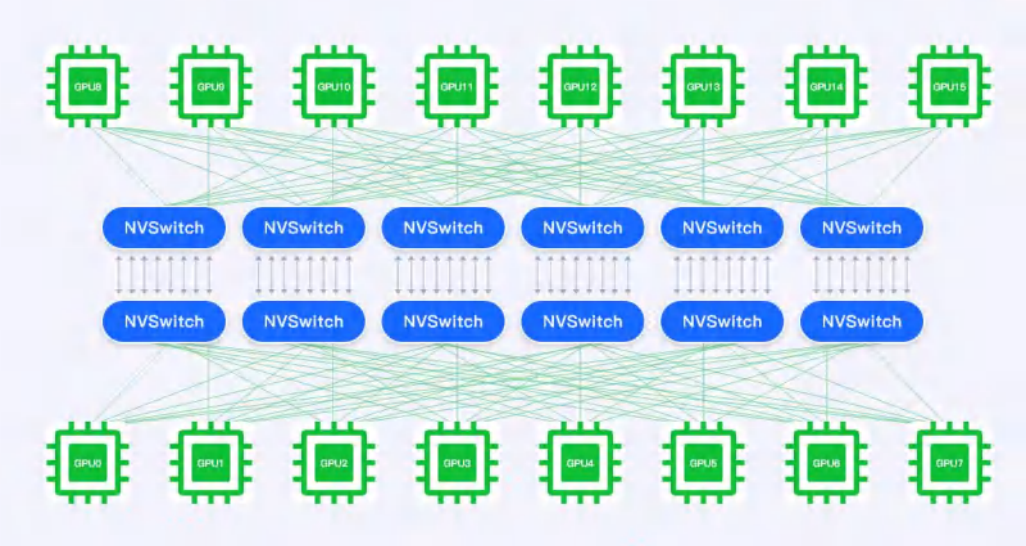

图 4 NVSwitch All-to-All 交换机架构

NVSwitch 的核心优势在于其能够构建全互联的 GPU 通信拓扑，所有连接到它的 GPU 都可以直接与其他任何 GPU 进行通信，无需经过中间节点或 CPU。这对于集合通信至关重要，它确保了所有 GPU 之间的数据交换都能以最高效率进行，避免了传统树形拓扑可能导致的收敛瓶颈。

NVSwitch 内部集成了大量的 NVLink 端口，每个端口都提供 NVLink 级别的高带宽和低延迟。例如，第三代 NVSwitch 系统架构中（如 DGX GB200），GPU 之间点对点的通信带宽可高达 900GB/s，这使得 NVSwitch 能够有效支撑大规模 AI 模型训练中海量的梯度同步和数据传输。除了 NVLink 端口，最新的 NVSwitch 还集成了对 400Gbps 以太网和 InfiniBand 连接的物理层支持。这意味着 NVSwitch 不仅可以作为 GPU 内部互联的桥梁，还可以作为连接外部网络的接口，从而实现 GPU 集群与数据中心网络的无缝融合，为构建更灵活、更强大的 AI 基础设施提供了可能。

（四）NVLink Fusion 开放互连技术

2025 年 5 月，英伟达推出了 NVLink Fusion 开放互连技术方案，允许第三方厂商（如高通、富士通等）的定制 CPU 或 AI 加速器通过 NVLink 协议与英伟达 GPU/CPU（如 Grace、Blackwell 系列）实现高速互联，支持单端口最高 900GB/s 的带宽，并集成 NVSwitch、Spectrum-X 交换机等组件构建机架级 AI 工厂。该技术通过开放生态策略，既保留了 NVLink 的低延迟优势（跨节点延迟<2 微秒），又支持异构计算（如 ASIC 与 GPU 协同）。

3.2.1.2 UALink

UALink（Ultra Accelerator Link）是 AMD、亚马逊 AWS、谷歌、英特尔、Meta、微软等公司共同发起的一项开放式互连标准，旨在为 AI 加速器集群提供高性能、高可靠、低成本的互连解决方案。

（一）UALink 的背景与目标

随着 AI 大模型参数量的爆炸式增长，对计算集群的规模和互连性能提出了前所未有的要求。英伟达的 NVLink 技术虽然性能卓越，但其专有性质使得其他厂商难以参与和创新，导致重复造轮子、生态碎片化等问题的出现。UALink 联盟的成立，正是为了解决这些痛点。

UALink 的目标是提供一个高性能、可扩展的互连标准，能够支持大规模 AI 加速器集群的构建，并满足不同 AI 工作负载的通信需求。它旨在实现以下关键特性：（1）开放性：UALink 是一个开放标准，任何厂商都可以参与其开发和采用，从而促进生态系统的繁荣和创新。（2）高性能：提供与 NVLink 相当甚至超越的带宽和低延迟，以满足 AI 训练和推理的极致性能需求。（3）可扩展性：支持连接数千甚至上万个 AI 加速器，构建超大规模的计算集群。（4）成本效益：通过标准化和开放性，降低互连解决方案的开发和部署成本。（5）灵活性：支持多种互连拓扑和通信模式，以适应不同的 AI 应用场景。

（二）UALink 架构与协议栈

UALink 1.0 规范是该标准的首个版本。UALink 1.0 支持每通道最高 200 GT/s 的数据传输速率，信令速率高达 212.5 GT/s，以满足以太网第 1 层前向纠错（FEC，Forward Error Correction）和额外第 1 层编码所需的带宽。UALink 通道可配置为单通道（x1）、双通道（x2）或四通道（x4）链路。四个通道组成一个站点（Station），在发送（TX）和接收（RX）方向各提供最高 800 Gbps 的带宽。这种灵活的配置使得加速器的数量和分配给每个加速器的带宽可以根据 AI 应用的需求进行灵活扩展。

UALink 交换机（ULS）可连接最多 1024 个加速器或端点，每个加速器被分配一个唯一的 10 位路由标识符，这使得 UALink 能够支持构建超大规模的 AI 加

速器集群。UALink 还支持将 Pod 进一步划分为多个虚拟 Pod。虚拟 Pod 是 Pod 内一个或多个加速器组成的逻辑组，组内加速器可以相互通信，但与 Pod 内其他加速器保持隔离。这种划分通过交换机端口子集的非重叠分配实现，提供了更灵活的资源管理和隔离能力。

UALink 采用物理层、数据链路层、事务层和协议层四层结构，自底向上深度融合性能优化与标准兼容性：

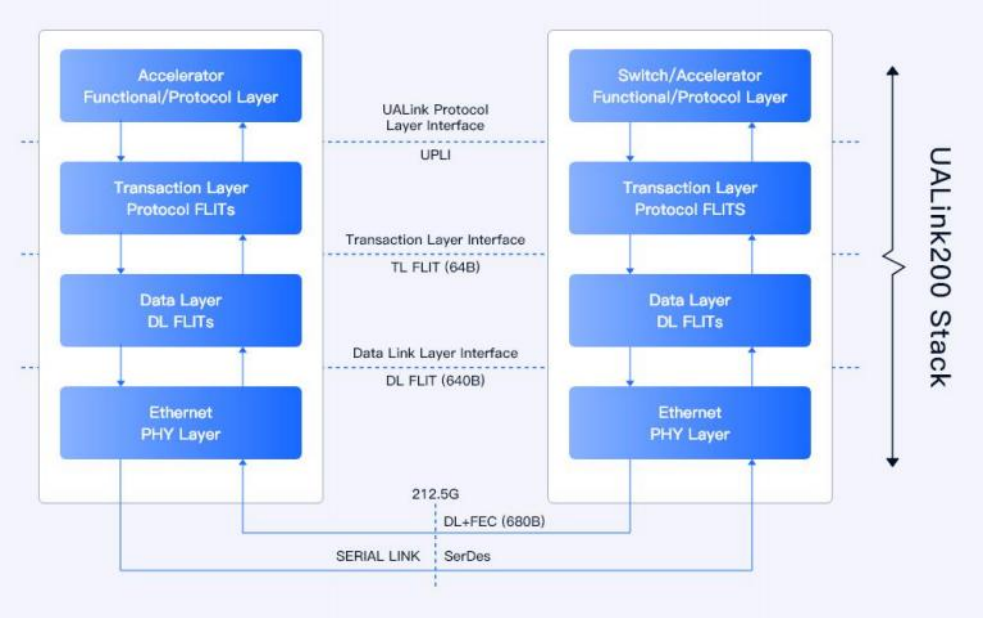

图 5 UALink 协议栈

物理层：UALink 物理层直接复用 IEEE 802.3dj 以太网标准，支持单通道 106.25 GT/s（低速模式）或 212.5 GT/s（高速模式），对应 100G 至 800G 带宽配置。物理编码子层（PCS）/物理介质连接接口（PMA）通过增强型码字交织模式，以优化 FEC 延迟。物理介质相关子层（PMD）、自动协商和链路训练（AN/LT）与 802.3 保持一致，未做修改。PCS 和 RS（协调层-PCA 与 DL 之间的接口）需要将 DL 信元与码字同步，以便将来自 DL 的 640 字节信元（生成一个 CRC）准确地装入一个 RS(544,514)码字。

数据链路层：数据链路层位于事务层与物理层之间，负责将来自事务层的 64 字节 Flit 打包聚合为 640 字节 Flit，供物理层使用。它还在链路伙伴之间提供消息服务，用于通告事务层速率、查询连接的链路伙伴的设备和端口 ID 等。消息服务还在链路伙伴之间提供一种 UART（通用异步接收器发送器）式的通信，主要用于固件（Firmware）通信。

事务层：事务层负责将来自 UALink 协议层接口（UPLI）的协议消息转换为事务层数据片（TL Flit），并将其传递给数据链路层。同时它也将从数据链路层接收到的 TL Flit 转换回 UALink 协议层接口上的协议消息。事务层支持流式地址缓存来压缩地址，以提高传输效率。

协议层：协议层是 UALink 协议栈的最上层，负责处理加速器之间的消息。UALink 是一种对称协议，在发送路径和接收路径中支持相同的消息集和信道。这些消息会经过 UALink 堆栈的多个功能层处理。

UALink 作为一项开放式的高性能互连标准，有望在 AI 加速器互连领域掀起一场技术革命，为构建未来超大规模 AI 计算基础设施奠定坚实基础。

3.2.2 新型互连技术

3.2.2.1 SUE

SUE（Scale Up Ethernet）是博通提出的一种新型互连框架，旨在将以太网的优势引入 AI 系统内部的 Scale Up 领域，实现 XPU（包括 GPU 等专用芯片）之间的高速、可靠、开放通信。

（一）SUE 的背景与目标

SUE 框架允许将 XPU 集群扩展至机架或多机架规模，以支持大规模数据集处理、深度神经网络训练及并行任务执行。其核心思想是以以太网为基础构建传

输层和数据链路层，直接在 XPU 间高效搬运内存事务。在部署模型上，SUE 支持单跳交换拓扑或直接互联的 Mesh 拓扑。每个 SUE 实例可灵活配置端口数(1/2/4 个），例如 800G 实例可拆分为 1×800G、2×400G 或 4×200G 端口，以适应交换机端口密度和冗余需求。单个 XPU 可集成多个 SUE 实例（如 8 或 16 个），通过多实例叠加实现超高带宽（例如 64 个 XPU 各配 12 个 800G SUE 时，任意 XPU 对间带宽达 9.6Tbps）。

（二）SUE 技术架构

技术架构上，SUE 采用类 AXI 的双工数据接口，通过虚拟通道（VC）将事务映射至不同流量类别，支持两种传输模式：严格有序模式（保障事务顺序）和无序模式（多端口负载均衡）。其协议栈分为三层：

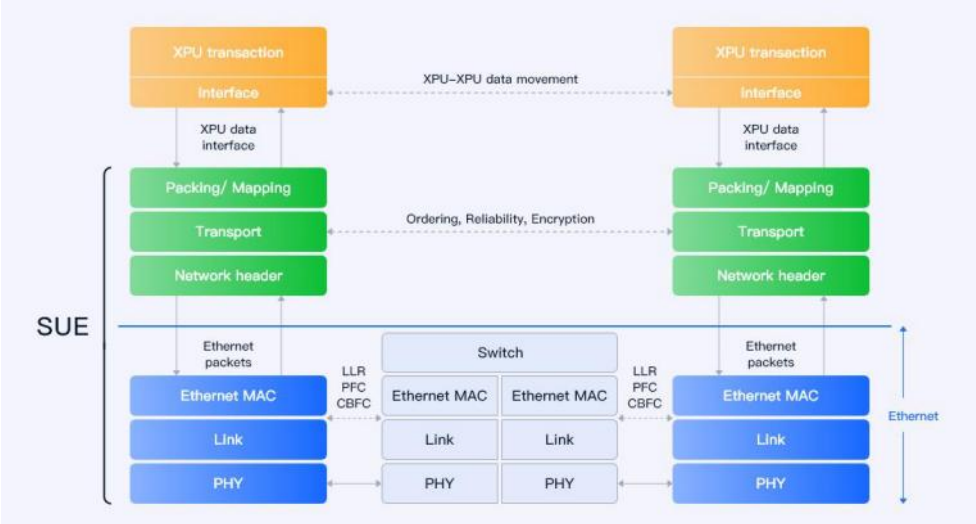

图 6 SUE 协议栈

映射打包层：将发往同一目标（XPU,VC）的事务聚合成最大 4096 字节的 SUE 协议数据单元（PDU，Protocol Data Unit）；

传输层：添加可靠性头部（RH），包含序列号（PSN）、虚拟通道（VC）及确认机制（RPSN），并附加 32 位 CRC 校验（R-CRC）。采用简化的 Go-Back-N

重传机制，通过 PFC/基于信用的流控（CBFC，Credit Based Flow Control）和链路层重传（LLR，Link Level Retry）实现无损网络；

网络层：支持标准以太网/IPv4/IPv6/UDP、优化的 AI 转发报头（AFH Gen1）及高度压缩的 AFH Gen2（仅 6-12 字节），以降低协议开销。

SUE 提供三类接口。XPU 命令接口：支持 FIFO 信用机制或 AXI4 总线，传输事务指令及数据（控制字段 144 位，含操作码、长度及目标 XPUID）；XPU 管理接口：基于 AXI 的寄存器配置通道；以太网接口：支持 200G/100G Serdes 速率，兼容 PFC/CBFC 流控及 LLR 重传，可动态切换故障链路。

SUE 实时监测各目标队列，在不引入额外延迟的前提下，将队列内多个事务动态打包成单个以太网帧发送（上限 2KB）。接收端通过 PSN 验证顺序，错误时触发 NACK 及重传。负载均衡由 XPU 层全局调度（跨多个 SUE 实例）和 SUE 内部调度（多端口无序模式）共同实现。SUE 要求端到端往返延迟（RTT）低于 2 微秒，单跳网络最多支持 1024 个 XPU。通过优化封装、无损流控及物理层技术（如空心光纤），10 米传输的单向延迟可控制在 520 纳秒内，满足 XPU 间内存事务的苛刻时延需求。

3.2.2.2 OISA

全向智感互联 (OISA, Omni-directional Intelligent Sensing Express Architecture) 是中国移动提出的开放 GPU 互连协议体系，旨在解决万亿参数大模型训练中的通信墙问题。

（一）OISA 的背景与目标

超大模型训练依赖 GPU 间频繁数据交互，通信开销导致集群有效算力无法随 GPU 数量线性增长，互联性能成为制约集群规模扩展和性能提升的瓶颈。OISA

旨在打造高效、智能、灵活且开放的 GPU 卡间互联体系，支持大模型训练、推理、高性能计算等数据密集型 AI 应用。

（二）OISA 协议架构

OISA 采用分层设计的协议结构，分别是事务层、数据层和物理层，允许各个层级专注于特定的功能实现，从而保障整个系统的优化和高效运行

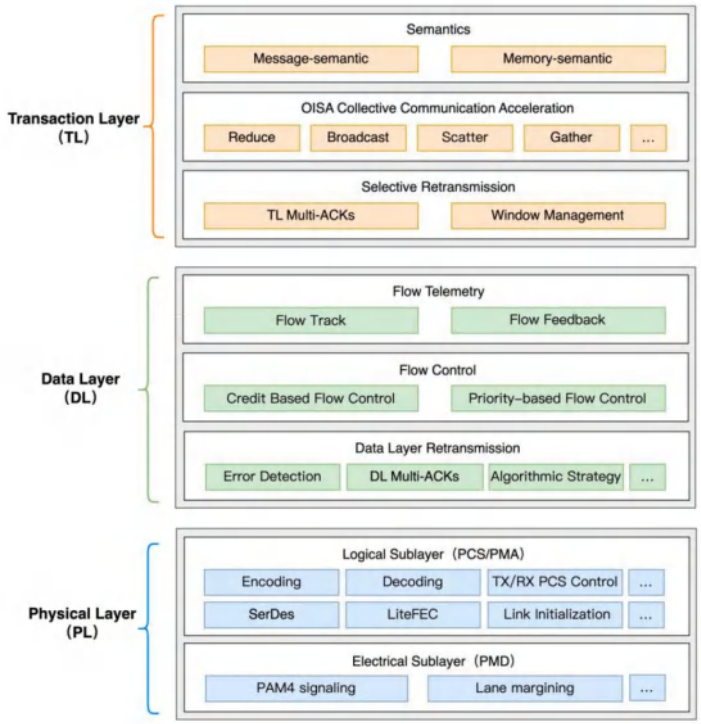

图 7 OISA 核心架构

事务层：最上层，负责封装数据，支持消息语义、内存语义、多语义三种模式，引入选择性重传（SR，Selective Repeat）机制，相比 Go-Back-N 协议在 Scale Up 场景效率更高，可精确识别并仅重传丢失或损坏数据包，提高传输效率。

数据层：介于事务层和物理层之间，定义了具有流量感知能力的报文结构，为实现链路资源和传输速率的动态调整提供依据。引入了 CBFC 和 PFC 两大流控技术，避免网络拥塞并确保数据的有序流动。此外还引入了数据层重传技术提升系统响应速度和可靠性。

物理层：最底层，包括逻辑子层和电气子层。逻辑子层定义了数据传输逻辑特性，负责编码和时序同步；电气子层负责将逻辑子层编码后的数据转换为电气信号，定义电气参数和接口，确保设备兼容性和连接正确性。

OISA 接口支持 AXI 总线接口（如 AXI Stream 用于高速数据传输，AXI Lite 用于控制信息传输），也可兼容 GPU 厂商自定义接口方案。

OISA 通过统一报文格式、多语义融合、多层次流控和重传、集合通信加速等关键技术，实现了高速、低时延、无损和高可靠的 GPU 通信。

3.2.2.3 ALS

在 2024 ODCC 开放数据中心大会上，阿里云联合信通院、AMD 等十余家业界伙伴发起 AI 网络互连开放生态 ALS（ALink System）。ALS 产业生态支持 UALink 协议，目标是解决 AI 智算超节点快速发展中面临的超高速、超大带宽的 Scale Up 技术难题。通过开放生态促进 AI 智算互连领域的技术创新和标准化。目前，ALS 已形成从协议到芯片、从硬件设备到软件平台的系统体系，在 ALS-D 数据面支持 UALink，在 ALS-M 管控面提供统一接口规范和管控软件平台。

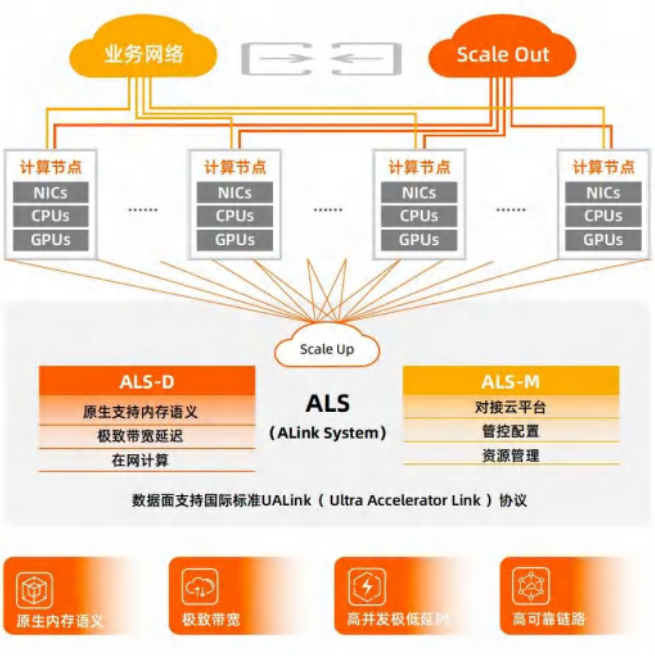

图 8 ALS 技术架构图

ALS-D 数据面：ALS-D 数据面互连采用 UALink 协议，这意味着它原生支持高性能内存语义访问和显存共享，并支持 Switch 组网模式。在性能上，ALS-D 具备超高带宽、超低时延能力，并增加了在网计算等特性。这使得 ALS-D 能够为 AI 应用提供强大的数据传输能力和计算加速。

ALS-M 管控面：ALS-M 管控面旨在为不同芯片提供标准化接入方案，符合规范的设备均可灵活接入应用方系统。无论是对开放生态还是厂商专有互连协议，ALS 都使用统一的软件接口。同时，ALS-M 为云计算等集群管理场景，提供单租、多租等灵活和弹性的配置能力，从而实现对 AI 集群资源的精细化管理和调度。

ALS 是 UALink 协议的积极实践者和推广者，通过将 UALink 协议应用于实际的 AI 基础设施中，验证了 UALink 的性能和可行性，并进一步丰富了 UALink 的生态系统。可以说，ALS 是 UALink 在 AI 智算超节点领域的一个重要应用和落地。

3.2.2.4 ETH+

高通量以太网（ETH+）协议是由中国科学院计算技术研究所、阿里云等超 40 家机构组成的高通量以太网联盟（ETH+ Consortium）发布的一种新型以太网协议。2024 年 9 月发布了 1.0 版本，基于 ETH+协议的网络协议 IP、开源网卡等硬件和系统也已经公开。ETH+旨在通过对以太网帧格式、链路层和物理层进行优化，以及结合 RDMA 在网计算（In-Network Computing）技术，显著提升以太网在 AI 智算网络中的性能，以应对 AI 时代对高效、稳定、可扩展网络的需求。

帧格式优化：ETH+协议通过优化以太网帧格式，有效提升了以太网帧的有效载荷比（可达 74%）。这意味着在相同带宽下，可以传输更多的有效数据，从

而大幅提高 AI 数据中心大量短数据报文的传输效率。对于 AI 训练中频繁出现的梯度同步等小包传输场景，这一优化尤为关键。

链路层与物理层重传：ETH+深度支持链路层和物理层的重传技术。传统的以太网在链路层通常不提供重传机制，一旦发生丢包，需要上层协议（如 TCP）进行重传，这将引入较大的延迟。ETH+在更低的层次引入重传机制，可以更快地恢复丢失的数据包，从而显著提升网络的语义可靠性，并降低端到端延迟。

RDMA 在网计算：ETH+基于 RDMA 技术，并进一步结合了在网计算能力。RDMA 允许应用程序直接访问远程内存，绕过 CPU 和操作系统，从而显著降低延迟和 CPU 开销。在网计算则允许网络设备在数据传输过程中执行简单的计算任务，例如集合通信中的聚合操作。通过这种结合，ETH+实现了集合通信性能30%以上的提升，有效解决了传统以太网在处理 AI 集合通信时的效率问题。

2025 年 8 月 14 日，高通量以太网联盟发布了 ETH+协议 1.1 版本。同时还推出了全量支持 ETH+特性的首款国产 400G 智能网卡芯片、首款国产 25.6T 交换芯片、ERack+/ORack+国产硅光芯片，以及首款 ETH+ 64 超节点等。

3.3 Scale Out（横向扩展）技术

3.3.1 InfiniBand 技术

在面向大模型训练的 AI 集群中，InfiniBand 凭借其高带宽、低延迟和原生 RDMA 能力，成为英伟达等厂商构建高性能训练网络的首选互连技术。

3.3.1.1 IB 架构与协议栈

集群结构上，InfiniBand 与 NVLink、NVSwitch 协同构建三层通信架构：

- 节点内通信：依赖 NVLink 或 NVSwitch，实现单节点内 GPU 间的高速互联；

- 节点间通信：通过 InfiniBand 网络实现跨服务器 GPU 通信，支撑分布式训练；

- 多节点组网：使用 InfiniBand 交换机（如 Quantum、Spectrum）构建 Fat-Tree 或 Dragonfly 拓扑，确保通信性能和扩展性。

在 DGX 系列中，一台服务器内部的 8 块 GPU 通过 NVSwitch 实现全互联，每台服务器配备多个 HDR 或 NDR InfiniBand 网卡，提供最高达 800 Gb/s 的通信带宽。服务器间通过 InfiniBand Spine-Leaf 网络结构连接，构建成标准的两层 Fat-Tree 或 Dragonfly+拓扑网络，以保障任意 GPU 对任意 GPU 的高带宽、低阻塞通信。

InfiniBand 的技术优势包括原生支持 RDMA 与零拷贝通信，避免 CPU 参与，大幅降低延迟；具备端到端通信可靠性机制（E2E Reliability），无需依赖软件协议重传；通过 Subnet Manager 管理网络拓扑、路径与服务等级，实现可编程与集中调度。

InfiniBand 采用分层协议栈设计，与 OSI 模型类似但针对高性能场景进行了优化：

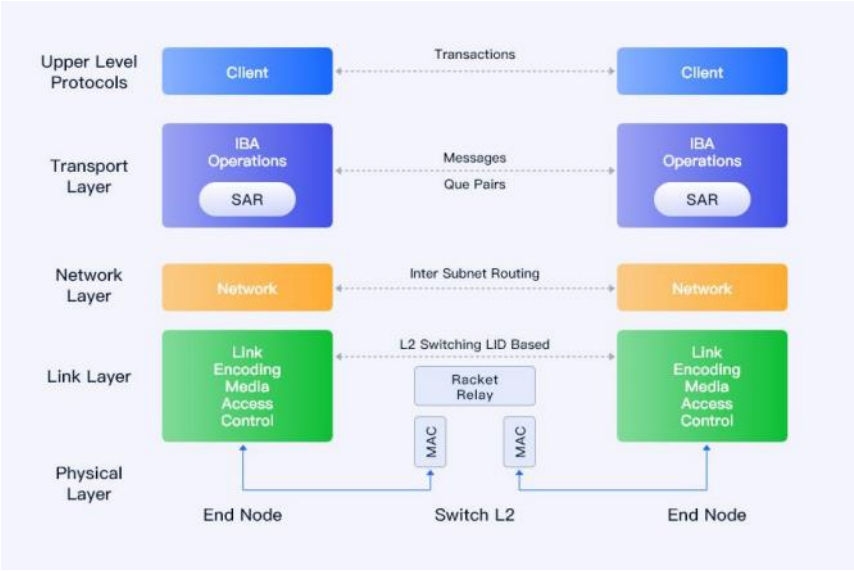

图 9 InfiniBand 协议栈

物理层：定义高速串行传输接口，支持多速率（HDR、NDR、XDR）和编码方式（NRZ、PAM4），确保高速稳定的信号传输。

链路层：负责数据帧的组装、流量控制、错误检测（CRC 校验）和虚拟通道管理。虚拟通道允许在同一物理链路上划分多个逻辑通道，实现流量隔离和优先级管理，避免热点和拥塞扩散。链路层还保证数据传输的可靠性和有序性，是 InfiniBand 实现端到端高性能通信的基础。

网络层：网络层实现数据包的路由与路径选择功能。InfiniBand 支持多种拓扑和多路径路由策略，包括静态路由和自适应路由，以提高网络负载均衡和容错能力。

传输层：传输层是协议栈的核心，提供多种通信服务类型：可靠连接（RC，Reliable Connection），面向连接，确保数据有序且无丢失（需确认），适合大多数训练同步通信；可靠数据报（RD，Reliable Datagram），无连接，支持多目标通信（需确认）；不可靠连接（UC，Unreliable Connection），面向连接，不保证交付（无需确认）；不可靠数据报（UD，Unreliable Datagram），无连接，不保证交付（无需确认）。

InfiniBand 协议栈从物理层到传输层全栈自研，内置 RDMA 与流控机制，保证硬件层面端到端的高效和可靠。

3.3.1.2 关键技术与带宽演进

InfiniBand 的核心优势之一是其底层通信协议中原生集成了 RDMA 机制，无需操作系统介入，即可通过 GPUDirect 技术实现节点之间的 GPU 直达通信，极大减少延迟与内存拷贝开销，是大规模训练任务通信效率的关键保障。

（1）RDMA 与 GPUDirect

InfiniBand 专为高性能通信设计，不依赖软件协议堆栈即可提供低延迟、零拷贝的数据交换路径。通过 GPUDirect RDMA 技术，网卡可以直接访问 GPU，跳过 CPU 和主内存，从而降低通信延迟至微秒级，并释放主机资源用于计算。

（2）拥塞控制与链路可靠性机制

InfiniBand 的另一个核心特性是其端到端的通信可靠性机制，协议层自带包序列管理、确认与重传机制，确保每个数据包都可靠送达，这一点使其在应对大规模 GPU 同步通信中的网络抖动和丢包问题上表现更优。同时它还支持服务等级与虚拟通道机制，用于多租户隔离和流量调度；自适应路径选择与 FEC，增强网络弹性与稳定性；以及 Subnet Manager，可控制拓扑路径、拥塞点规避与 QoS 策略。

（3）带宽演进

InfiniBand 的物理层带宽已历经数代跃升，从早期的 10Gb/s 到最新单端口 800 Gb/s：

表 2 InfiniBand 演进

	单通道速率（Gb/s）	四通道总速率	编码方式	产品代表
FDR	14	56 Gb/s	NRZ	–
EDR	25	100 Gb/s	NRZ	–
HDR	50	200 Gb/s	NRZ	DGX A100
NDR	100	400 Gb/s	PAM4	DGX H100
XDR	200	800 Gb/s	PAM4	Blackwell

3.3.1.3 InfiniBand 与以太网

在实际部署中，InfiniBand 与以太网形成互补关系。二者的核心差异主要体现在协议架构、性能稳定性与管理方式等方面：

表 3 InfiniBand 与以太网

维度	InfiniBand	Ethernet + RoCEv2
协议栈	原生 RDMA、硬件可靠	基于 UDP/IP 的以太网协议栈
延迟	亚微秒级（典型值 600ns~1μs）	微秒级（典型值 400ns~2μs）
带宽	当前主流 400Gbps，支持 NDR 800Gbps	依赖以太网物理层，支持 800GbE
拥塞控制	基于信用的流控 + 硬件级重传（零丢包）	依赖 PFC、ECN、DCQCN 等无损机制
异构环境	封闭生态，需专用网卡 / 交换机	开放标准，支持多厂商设备混合部署
硬件成本	高（专用设备溢价 30%~50%）	低（复用以太网交换机，成本约 IB 的 1/3）
运维复杂度	低（厂商集成方案，开箱即用）	高（需调优 PFC/ECN 避免死锁和风暴）

在以太网生态逐步引入 RoCEv2、CXL 互联协议的背景下，也出现了 InfiniBand 与以太网融合的趋势，例如英伟达 Quantum-2 同时支持 IB 与 Ethernet。

3.3.2 RoCEv2 技术

RoCE 是一种通过以太网网络进行远程直接内存访问的网络协议。和 InfiniBand 协议一样，RoCE 协议通过发送方直接将数据写入接收方的内存中，无需经过接收方的操作系统内核，从而实现高带宽，低延迟的数据传输。RoCE 协议可以无缝地集成到已有的以太网网络架构中，而无需更换交换设备，因此得到了产业界的广泛支持。

3.3.2.1 高性能 RDMA 技术

目前 RoCE 的主流版本有两个：RoCEv1 和 RoCEv2，分别由 IBTA 于 2010 年和 2014 年推出。RoCEv1 在网络层和传输层复用了 InfiniBand 协议，仅在链路层使用以太网协议，允许在同一以太网广播域中的任意两台主机之间进行通信。而 RoCEv2 则将协议扩展到了网络层，这意味着 RoCEv2 数据包可以被现有以太网协议设备路由，进一步提高了与现有设备的兼容性和适用范围。因此，在现代智算网络中，RoCEv2 取代了 RoCEv1，成为和 InfiniBand 一样的主流 RDMA 网络协议技术。

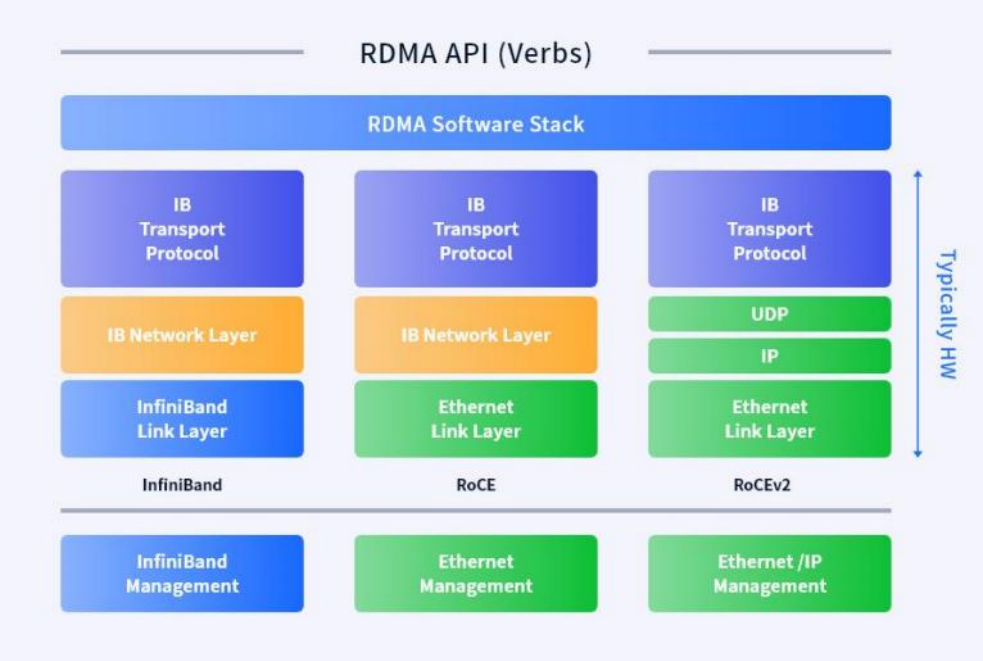

图 10 InfiniBand、RoCE、RoCEv2 对比

RoCEv2 通过以下特性实现高性能数据传输：

零拷贝数据传输：应用程序可直接访问远程内存，绕过操作系统内核，减少数据拷贝次数（从传统 TCP/IP 的 4 次拷贝降至 1 次），显著降低 CPU 开销。在 100Gbps 网络下，RoCEv2 的端到端延迟可低至 1 微秒以内，优于 TCP/IP 的 10-20 微秒。

高效协议栈：基于 UDP 封装，省去 TCP 的连接建立、拥塞控制等开销，保持以太网帧格式不变，便于网络设备处理。支持大规模并发连接，单网卡可支持数百万个活跃队列对，满足微服务、容器化环境的高并发需求。

内核旁路：RoCEv2 通过用户态驱动直接操作网卡硬件，避免了内核上下文切换。每次通信可节省约 2000 个 CPU 时钟周期，这对于高频交易等场景至关重要。

硬件加速：通过网卡中的专用 ASIC 芯片实现 RDMA 操作的硬件卸载，进一步提升性能。支持多队列并行处理，利用现代 CPU 多核优势，实现线性扩展的吞吐量。

3.3.2.2 跨网段路由技术

RoCEv2 相较于其前身 RoCEv1 的一个显著优势在于其支持跨网段路由。RoCEv1 仅限于在二层网络（同一广播域）内进行 RDMA 通信，而 RoCEv2 通过将 RDMA 协议封装在 UDP/IP 报文中，使其能够在三层网络（IP 网络）中进行路由，从而实现了跨网段的 RDMA 通信。

RoCEv2 的数据包在以太网帧的基础上，增加了 IP 头和 UDP 头。这种封装方式使得 RoCEv2 数据包可以像普通的 IP 数据包一样，在标准 IP 路由器和交换机之间进行转发。具体来说，RoCEv2 将 RDMA 传输层协议的数据封装到 UDP 数据报中，然后 UDP 数据报再封装到 IP 数据包中。IP 头包含了源 IP 地址和目的 IP 地址，使得数据包可以在不同的 IP 子网之间进行路由。UDP 作为传输层协议，提供了端口号，用于区分不同的 RDMA 连接。虽然 UDP 本身是无连接和不可靠的，但 RoCEv2 在 UDP 之上实现了可靠传输机制，确保了 RDMA 通信的可靠性。

RoCEv2 突破了传统二层网络的限制，让 AI 集群可以部署在更大的规模和更复杂的网络拓扑中，支持数千甚至上万个节点的互联。此外，RoCEv2 允许 AI 计算节点和存储节点分布在不同的子网中，提高了网络部署的灵活性和资源利用率。RoCEv2 可以在现有的标准以太网基础设施上部署，无需对网络设备进行大规模升级或更换，降低了部署成本。

然而，为了确保 RoCEv2 在跨网段路由时的性能，仍然需要关注底层以太网的无损特性。虽然 IP 路由提供了跨网段的能力，但如果底层以太网存在丢包或

拥塞，仍然会影响 RDMA 的性能，因此在部署 RoCEv2 跨网段网络时通常需要配合无损以太网技术（如 PFC、ECN 等）来保证端到端的无损传输。

3.3.2.3 拥塞管理与流控机制

传统的以太网是“尽力而为”的传输模式，在拥塞时会丢弃数据包，这对于 RDMA 来说是不可接受的。因此，RoCEv2 依赖于无损以太网技术来确保数据传输的可靠性。RoCEv2 的拥塞管理和流控机制主要通过以下技术协同工作：

PFC：PFC 是 IEEE 802.1Qbb 标准定义的一种链路层流控机制，也被称为逐跳流控。它允许网络设备根据数据包的优先级对流量进行暂停。当交换机某个端口的缓存达到预设阈值时，它会向发送方发送一个暂停帧（Pause Frame），通知发送方停止发送特定优先级的数据，直到缓存压力缓解。PFC 能够有效防止链路层丢包，是构建无损以太网的基础。

ECN：ECN 是一种网络层机制，允许网络设备在检测到即将发生拥塞时，通过在 IP 报头中设置 ECN 标记来通知发送方。当交换机检测到队列深度超过某个阈值时，它会在数据包的 IP 头中设置 ECN 位，而不是直接丢弃数据包。接收方收到带有 ECN 标记的数据包后，会将其转发给发送方，发送方根据 ECN 标记降低发送速率，从而避免拥塞的发生。

数据中心量化拥塞通知（DCQCN，Data Center Quantized Congestion Notification）：DCQCN 是 RoCEv2 中一种端到端的拥塞控制算法，它结合了 ECN 和基于速率的拥塞控制。DCQCN 通过以下步骤工作：（1）拥塞标记：交换机在检测到拥塞时，通过 ECN 标记数据包。（2）拥塞通知包（CNP，Congestion Notification Packet）：接收方收到带有 ECN 标记的数据包后，会生成一个 CNP 并发送给发送方。（3）速率调整：发送方收到 CNP 后，会根据 CNP 中的信息

（如拥塞程度）动态调整其发送速率。DCQCN 采用了一种量化的速率调整机制，使得发送方能够快速响应拥塞并恢复到最佳发送速率。

这些机制协同工作，确保了 RoCEv2 在以太网环境下的高性能和无损传输。PFC 提供了链路层的无损保障，ECN 提供了网络层的拥塞预警，而 DCQCN 则实现了端到端的拥塞控制和速率调整。通过这些技术，RoCEv2 能够为 AI 集群提供稳定、高效的 RDMA 通信，满足其对网络性能的严苛要求。

3.3.3 UEC 传输协议

超以太网联盟(UEC，Ultra Ethernet Consortium)的成立标志着以太网技术演进的一个重要转折点。UEC 联盟由 AMD、Arista、Broadcom、Cisco、HPE、Intel、Meta 和 Microsoft 等行业巨头于 2023 年联合发起，旨在解决传统以太网在 AI 和 HPC 场景下面临的性能瓶颈。2025 年 6 月，UEC 规范 1.0 版本正式发布，这被视为以太网向新一代数据密集型基础设施演进的关键一步。

3.3.3.1 UEC 传输协议

UEC 协议代表了以太网技术在高性能计算和人工智能时代的一次重大革新。作为对传统以太网和 InfiniBand 技术的突破性升级，UEC 协议从底层物理层到上层应用接口进行了全方位重构，旨在满足现代 AI/HPC 工作负载对网络的严苛需求。

UEC1.0 提供的两个重要特性 LLR 和 CBFC 是用于支撑内存语义实现的重要基石。LLR 的工作机制是在链路层实现点到点的可靠传输。发送侧为每一个传送的 L2 帧分配一个序列号。该序列号嵌入在 L2 帧的前导码中传输。接收侧负责校验收到的 L2 帧携带依次增加的序列号，并周期性地发送 ACK 以向链路伙伴指示最近成功接收的 L2 帧的序列号。如果接收的 L2 帧 FCS 校验错误或者序列号非预期，接收侧发送 NACK 请求发送端重传。发送侧维持一个重传缓冲，重传

的机制采用 Go-back-N。在 Scale Up 网络中采用 LLR 的核心诉求在于本地发生的错误就近发现，就近恢复。相较而言，如果不采用 LLR 机制，发生丢包/错包之后依靠端到端的选择性重传机制来恢复，那么就会给内存语义访问带来极大的延时增加。

CBFC 通过信用机制控制链路上的 L2 帧传输，发送方以信用为单位跟踪接收方可用的缓冲区空间，并且只有当接收方有足够的缓冲区空间（以信用为单位）时，调度器才允许调度数据包从无损 VC 队列进行传输。CBFC 消息用于将信用从接收方返回给发送方。接收方的信用生成取决于接收方端口缓冲区的可用性。发送侧以相对低的频度向接收侧发送已使用的信用计数器更新，此更新的目的是重新捕获由于链路错误（例如无法纠正的 FEC 错误或数据包 CRC 错误）导致的数据包丢失而可能“泄露”的信用。此外，CBFC 支持多达 32 个 VC，可以为每个 VC 独立配置优先级和信用分配。在 Scale Up 网络中采用 CBFC 的核心诉求在于实现无损传输，避免由于接收侧没有足够的缓冲导致丢包。这种情形的丢包只能依靠端到端的选择性重传机制来实现，同样会给内存语义访问带来极大的延时增加。

从线路传输效率的角度来说，Scale Up 网络中内存语义的操作主要是以 Cache Line(256B/128B)为颗粒度的内存访问，传统基于以太的协议栈报文封装势必给内存语义的操作带来巨大的开销。UEC 的多个相关技术工作组已经或正在开展相关的研究工作。以 Link WG 为例，工作组早已启动优化以太网 L2 header 的研究，相关的提案如统一转发头(UFH，Unified Forwarding Header)旨在提供与标准 Ethernet 无缝的互联互通，同时通过压缩提供更好的线路传输效率，而并非复制或者取代标准 L2+L3 header 所提供的功能。UFH 提案起初是为 Scale Out 网络设计的，但是也同样适用于 Scale Up 网络。

表 4 UEC 协议栈

协议栈分层		封装开销
Ethernet Link Layer	IPG	12B
	Preamble	8B
	L2 Header	14B
	FCS	4B
Networking	IP	20B
	UDP	8B
UET Transport	PDS	12B
	SES	4B

3.3.3.2 UEC 架构与协议栈

UEC 1.0 采用模块化分层架构，通过软件层开放架构接口 LibFabric、传输层 UET 的多维创新、网络层的数据包修剪，以及链路层的性能增强等关键技术实现了全栈优化。

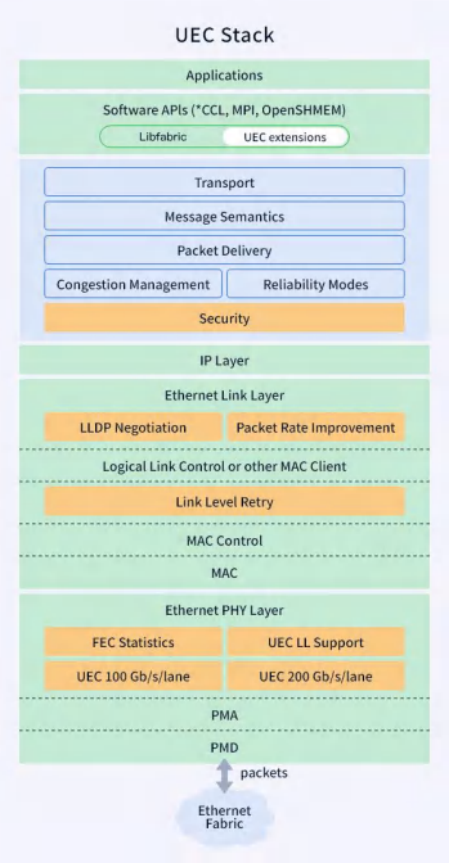

图 11 UEC1.0 协议栈

软件层：LibFabric 2.0 API 是 UEC 软件栈的核心构建模块，定义了一套面向高性能并行和分布式应用程序的通信接口。LibFabric 的主要目标是提供统一接口，让开发者能够方便构建应用，而无需关心底层具体的传输协议和硬件细节。LibFabric 支持 TensorFlow、PyTorch 等 AI 框架及 NCCL、MPI 等 HPC 库，通过 JobID 实现多租户隔离。单个 NIC 可承载多个 Fabric End Point（FEP），每个 FEP 仅属于一个 Job 并与同 JobID 的 FEP 交互。

传输层：传输层是 UEC 规范的核心创新所在。语义子层解析应用请求（如 AllReduce），定义消息语义与操作类型（发送/读取/原子操作），实现计算指令的优先级区分；分组交付子层提供可靠无序（RUD,Reliable Unordered Delivery）、可靠有序（ROD、Reliable Ordered Delivery）、可靠无序幂等（RUDI，Reliable Unordered Delivery For Idempotent）、不可靠无序（UUD，Unreliable Unordered Delivery）四种交付模式，适应不同应用场景；拥塞管理子层支持动态窗口调整与流量分类。传输安全子层定义 AES-GCM 加密算法与密钥管理机制，保障数据传输安全。

网络层：支持包修剪（Packet Trimming），允许交换机截断有争议的数据包，修改截断数据包的 DSCP 字段，并将其作为拥塞信号转发到目的地，为上层协议提供更多拥塞信息，以确保快速重传丢失的数据包。

链路层：UEC 链路层最大的变化是引入了 LLR 协议，它可以让以太网不依赖 PFC，实现无损传输，本地化错误恢复将重传延迟压缩至 1 微秒内，相比端到端重传效率提升 100 倍。

物理层：UEC1.0 规范下的物理层与传统以太网完全兼容，支持每通道 100Gbps 和 200Gbps 速率，并在此基础上实现 800Gbps 和更高的端口速率。

3.4 前沿突破技术

3.4.1 确定性广域网技术

AI 训练与推理对网络性能要求严苛，大模型分布式训练需跨数据中心协同，传统网络难以保障。确定性广域网技术能提供高可靠、低延迟、高带宽的传输服务，显著提升网络在带宽、时延、抖动、丢包等维度的指标，让网络从“尽力而为”转变为“确保所需”，为 AI 业务提供确定性服务质量，成为推动 AI 持续创新发展的关键支撑技术。

3.4.1.1 DetNet

随着工业互联网、自动驾驶、远程医疗以及 AI 训练等对网络时延、抖动和丢包率有严格要求的应用场景的兴起，传统尽力而为的 IP 网络已无法满足其确定性传输的需求。确定性网络应运而生，旨在为这些高实时性业务流提供可预测、可规划的网络传输服务，确保极低的丢包率和确定的端到端传输时延及抖动。

DetNet 是互联网工程工作小组（IETF，Internet Engineering Task Force）提出的一个标准体系，其目标是在 IP 层和以太网层提供确定性服务。DetNet 架构的核心思想是通过对网络数据转发行为的精确控制，实现可预期的性能。其核心技术包括：

资源预留与调度：通过划分转发时隙、资源预留（链路带宽预留、节点缓存预留等）和包抢占实现超低延迟和零拥塞损失，不仅限制端到端时延上界，还控制时延下界，实现更低时延抖动。允许在每个调度周期内，将确定性流的空闲资源调度给非确定性流使用，实现弹性调度。

可靠传输：采用包复制和冗余消除技术，在入口边缘节点复制数据包并通过多个路径发送，网络边缘节点进行冗余副本消除和原始数据包还原，确保单个随机事件或设备故障不会导致数据包丢失。支持多发选收功能，首节点复制报文在多条链路上同时发送，尾节点消除冗余副本并重新排序，提升网络可靠性。

路径控制：通过特定协议或集中控制单元计算确定性业务流的最佳路径，并依靠冗余路径保证个别链路故障时业务不中断，消除协议收敛时间的影响，保持转发路径稳定。

作为新型 IP 层确定性网络技术，DetNet 通过精准控制时延、抖动和丢包率，为 AI 应用提供确定性服务质量保障。

3.4.1.2 DIP

确定性 IP（DIP，Deterministic IP）网络是一种在 IP 网络中为端到端报文转发提供确定性时延和抖动的技术。与 DetNet 在 IP 层提供调度保障不同，DIP 技术更侧重于在传统 IP 网络的基础上，通过创新的机制消除因数据突发带来的转发抖动，从而充分保障网络报文的确定性传输。DIP 技术尤其适用于对时延和抖动有严格要求的工业控制、机器视觉、远程协作等 AI 应用场景。DIP 技术通常基于 SRv6（Segment Routing over IPv6）网络部署，并结合了边缘整形、门控机制、周期调度等关键技术。

边缘整形：在网络入口设备上通过令牌桶算法对突发流量进行周期性整形，将不规则的报文流转换为固定周期的数据流。令牌桶以恒定速率生成令牌，当报文到达时消耗对应数量的令牌。若令牌不足，报文将被缓存或丢弃，从而强制流量符合预设的平均速率和突发容量。

门控调度：网络节点采用定时轮循机制，将时间划分为固定长度的调度周期。每个周期内，设备仅在特定时段开放门控队列发送报文。例如，周期开始时清空队列并发送报文，其他时间队列处于关闭状态，强制报文按周期节奏传输。

周期映射：相邻节点通过周期映射协议自动协商发送周期，确保报文在转发过程中不会因时间差导致乱序。设备通过 DIP 学习报文（携带时延差值信息）动态计算本地转发周期。

DIP 技术的目标是实现微秒级的确定性时延，这对于需要高精度同步和实时响应的 AI 应用至关重要。例如在高度自动化的工业制造场景中，DIP 网络能够保障工业控制信号的超低时延，从而实现生产过程的精确控制和高效协同。

3.4.1.3 CSQF

指定周期排队转发（CSQF，Cycle Specified Queuing and Forwarding）是一种基于多周期的多队列循环调度机制，它是对标准 IEEE 802.1Qch 的循环排队转发（CQF，Cyclic Queuing and Forwarding）方法的一种扩展，主要为解决长延迟或跨域同步不精确的问题。作为一种面向多跳、长链路场景的确定性转发机制，CSQF 被认为是时间敏感网络（TSN，Time-Sensitive Networking）和 DetNet 等现有机制在可扩展性和跨域能力上的关键补充。

在 TSN 标准体系中，CQF 机制已经可用于在本地局域网中实现准确定时转发。但 CQF 存在两大局限：（1）对时钟同步精度要求极高，否则周期间包转发会错位；（2）仅支持两队列模型，当网络存在多跳或传播时延增加时，容易发生拥堵或传输周期错位。为了解决这些问题，CSQF 在 CQF 基础上扩展了三大能力：（1）支持指定周期排队与转发；（2）引入容忍队列机制；（3）实现端到端周期一致性传播。

CSQF 的核心思想是，在网络设备的出端口上划分周期，一个周期内对确定性业务流进行统一调度，使其在确定的时间片内进行转发。与传统的尽力而为转发机制不同，CSQF 通过精确控制数据包的转发时机，确保数据流能够按照预定的时间表穿越网络，从而实现端到端的确定性传输。

周期划分与时间同步：CSQF 引入全网络时间同步，交换机和网卡将时间分段为一致的周期，每个周期持续固定时长。

指定周期排队：每个流被分配一个目标周期。数据包根据其周期标签入相应队列，从而确保只有在指定周期才会被转发。与 CQF 使用两个队列不同，CSQF 至少使用三队列结构：其中偶队列（Even Queue）和奇队列（Odd Queue）交替进行接收和发送操作，容忍队列（Tolerating Queue）用于处理由于同步偏差或传输延迟导致错过预定周期到达的包。

精准周期转发：在每个周期开始时，所有交换节点根据周期标签进行统一的定时转发，逐跳保持端到端的延迟可预测性。

CSQF 技术通过精密的调度和排队机制，有效地解决了传统网络中时延和抖动不确定的问题。在实时控制、传感器数据采集和处理等 AI 应用中，CSQF 能够提供高精度、低时延的网络传输保障，为 AI 系统的稳定运行和高效决策提供坚实的基础。

3.4.1.4 长距 RDMA 技术

传统 RDMA 主要应用于数据中心内部环境，覆盖范围通常不超过 10 公里，典型应用包括高性能计算和 AI 训练。长距 RDMA 是指将 RDMA 协议从数据中心内部延伸至广域网环境，覆盖数百至上千公里的物理距离，实现跨地域算力中心之间的高吞吐、低时延、零丢包级别的数据传输。长距 RDMA 继承了 RDMA 的所有优势，适用于异地算力协同、海量数据搬迁等场景。

然而，将 RDMA 扩展到广域范围面临三大挑战：第一，传输时延变长，ACK 响应延迟。广域链路会导致 RTT 大幅提升，使得发送端 RDMA 网卡缓存快速填满，从而限制吞吐。第二，链路丢包影响严重。RDMA 使用 Go Back N 重传机制，对丢包高度敏感。一旦出现丢包，需要重传该点之后的所有数据包，吞吐显著下降。第三，网络流控不适配。DCN 交换机缓存不足，难以支撑长距 RTT 带来的

流控回馈滞后，导致拥塞无法及时应对。因此必须对架构、技术与协议等多方面进行优化和改进，提高 RDMA 跨广域传输吞吐率。具体措施包括：

全光网络承载：构建算间全光高速平面，将 DCN 网络的 Spine/Leaf 节点直连光传送网络（OTN，Optical Transport Network）光传输设备，OTN 设备基于物理层参数数据与端侧业务参数协同，实现高吞吐长距离传输。采用 Mesh 化、立体化拓扑进行组网，全面部署光交叉连接（OXC，Optical Cross-Connect），通过联动 OTN 实现光电协同高效调度。

协议优化：针对 RDMA 在广域网中的传输效率问题，优化 RDMA 协议参数，如调整 QP 数量和块大小，以确保最大吞吐率，适应不同距离和带宽条件下的传输需求。

长距 RDMA 是高性能广域算力互联的重要技术基础，它并不是对传统 RoCE 或 InfiniBand 的简单延伸，而是一种从网络架构到端侧协议栈全面协同优化的系统级能力。通过构建全光无损传输网络、优化协议参数、实现光电协同调度与端网信息共享，RDMA 能够实现百公里级范围的无损、高通量传输能力。

3.4.1.5 确定性光电融合技术

当前，国内外大厂正联合推动 IP 支持 ZR+，其背后的驱动力在于 OTN 芯片的发展跟不上 IP 流量的迅猛增长，导致网络运力与算力、数据要素发展不平衡。与此同时，AI 时代的众多应用场景迫切需要一个具备确定性的光电融合网络，以满足算网一体、一网多用的需求，实现时延、带宽的全颗粒灵活调配，按需定制。因此，确定性光电融合路由技术应运而生，成为解决当前网络挑战的关键技术之一。紫金山实验室和江苏未来网络集团在该领域取得了重大技术突破，攻克了“IP+光 ZR++技术”、“确定性网络技术”以及“CENI 大网操作系统光电融

合调度技术”三大关键技术，实现了光、电、算深度协同，重构广域网络的架构与控制逻辑。

（一）光电融合

光电融合的核心原理是将复杂的多层网络（WDM、OTN 和分组传输层）统一为一个易于控制的层。OTN 的电层功能被 IP 层吸收，路由器直接搭载 400G ZR+可插拔相干光模块，跳过传统 OTN 转发器，直接从路由器端口提供相干波长，让 IP 路由器具备光传输能力，实现 IP 层与光层的紧密协作。这种直接在路由器上发挥光传输功能的做法减少了中间转换环节，提高了传输效率。

ZR+采用 QSFP-DD 封装，符合 OIF（光互联论坛）协议，任何支持开放标准的路由器均可搭载，避免了厂商锁定；通过优化光电器件设计和信号处理算法，显著提高了光接收功率范围；引入增强型 oFEC 和优化算法，进一步提升了光接收的可靠性和传输距离，可实现超远距离无电中继无损传输；跳过 OTN 层，路由器直驱光层，使得综合成本大幅降低，同时攻克了长距传输、确定性时延与低成本三大难题。

（二）确定性网络

确定性网络作为核心支撑技术，其作用是通过构建可预测、可规划的网络传输环境，解决跨区域算力协同中的传输难题，满足工业互联网、远程医疗、AI 训练等场景对低时延、高可靠性的严苛需求。通过采用 DIP、CSQF 等技术，将网络传输划分为周期性时间片，确保确定性业务在固定时间片内传输，实现几乎零丢包、微秒级时延抖动、传输效率大于 90%的高质量网络传输能力。

（三）大网操作系统光电融合技术

CENI 大网操作系统支持用户根据业务需求（如带宽、时延、可靠性）动态调整网络资源，实现“分钟级”网络切片配置，满足不同场景的差异化需求。大

网操作系统光电融合技术通过 SDN 控制器集群，将 IP 路由器、数据中心交换机等异构设备纳入统一资源池，实现跨层资源全局可视与动态调度。同时通过智能编程技术，实现更合理的网络资源规划、业务的自动化开通和可管可控。

在实际应用中，确定性光电融合路由技术展现出了卓越的性能。在 CENI 现网的测试中，实现了沿江 2000 公里的远距传输，且在 400G 满负载的情况下，保持零丢包的无损传输。在运营成本上，通过全颗粒切片和“IP+光”融合调度，全网带宽利用率达到 90%以上，丢包率小于十万分之一，有效减少了网络维护和优化的成本。同时，设备的小型化和功耗的大幅降低，从 400G 功耗约 330W 降到约 30W，进一步节省了能源消耗和空间占用。

3.4.2 超节点计算架构

超节点（SuperPod）是近年来为应对 AI 大模型训练与推理需求而发展起来的新型算力基础设施架构，它通过高速互连技术将大量计算单元（如 GPU、TPU、NPU 等）紧密集成，构建一个高带宽域（HBD，High Bandwidth Domain）。具体来说，超节点是指在一个物理机柜或一组紧密耦合的计算单元内部，通过高密度集成计算单元和专用的高速互联技术，实现近似单机性能的超大规模并行计算系统。它旨在突破传统服务器内部以及服务器之间通过 PCIe 或标准以太网互联的带宽和延迟瓶颈，将数十甚至数百个加速器紧密连接，形成一个逻辑上的超大服务器，以支持张量并行、专家并行等对内部通信要求极高的并行计算任务。

超节点计算架构的关键特征包括：

- 高密度算力集成：在有限空间内集成大量 GPU 或其他 AI 加速器，提供极致的计算密度。
- 高速互联：采用 NVLink、InfiniBand 等高速互联技术，实现 GPU 之间以及 GPU 与网络之间的高带宽、低延迟通信，消除数据传输瓶颈。

- 算力与网络深度融合：网络不再仅仅是数据的传输通道，而是与计算紧密结合，实现网络感知计算、网络融合计算，甚至计算重塑网络。例如，在超节点内部，通过引入节点内交换芯片，增强卡间 P2P 带宽，有效提升节点内网络传输效率。

- 统一资源管理与调度：实现计算、存储和网络资源的统一纳管和融合路由调度，提升资源利用率和管理效率。

当前业界典型的超节点方案包括：

（一）英伟达 DGX SuperPOD(以 NVL72 为例)

英伟达作为 AI 加速领域的领导者，其 DGX SuperPOD 系列是业界广泛采用的 AI 超级计算平台。其中，GB200 NVL72 SuperNode 是其最新的代表性产品之一。GB200 NVL72 SuperNode 将 36 个 Grace CPU 和 72 个 Blackwell GPU 集成到一个液冷机柜中。它采用“GPU-GPU NVLink ScaleUp + Node-Node RDMA ScaleOut”的互联方式。

图 12 英伟达 DGX SuperPOD

- Compute Tray：整个系统包括 18 个 Compute Tray，每个 Compute Tray 包含 2 个 GB200 超级芯片，每个 GB200 超级芯片又包含 2 个 Blackwell B200 GPU 和 1 颗 Grace CPU，整个机柜共 72 个 B200 GPU 和 36 个 Grace CPU。通过 NVLink

和 NVLink-C2C 技术，实现 GPU 之间以及 GPU 与 CPU 之间的高速内存共享和数据传输。单个 Compute Tray 提供 7.2TB/s（单向 28.8Tb/s）带宽，NVL72 整机柜的 Compute Tray 提供 129.6TB/S 的 NVLink 带宽。

- Switch Tray：共包含 9 个 Switch Tray，每个 Switch Tray 内置 2 颗 NVSwitch 芯片，整个机柜提供 18 个 NVLink Switch 芯片。整机柜后部通过线缆将 Compute Tray 和 Switch Tray 进行互联。单个 Switch Tray 提供 14.4TB/s（单向 57.6Tb/s）带宽，NVL72 整机柜的 Switch Tray 提供 129.6TB/s 的 NVLink 带宽。这样超节点整机柜 Compute Tray 的 GPU 和 Switch Tray 的交换芯片之间就可以实现全连接。

- Scale Up：NVL72 内部采用 NVLink5 和 NVSwitch 构建 Scale Up 网络，提供极高的带宽 (每个 Compute Tray 通过 NVLink/NVSwitch 具有 7.2TB/s 的 ScaleUp 连接带宽）和超低时延（铜电缆连接节省了光模块引入的时延）。所有 GPU 可以访问整个超节点其他 GPU 的 HBM 内存和 Grace CPU 的 DDR 内存，实现统一内存空间。

- Scale Out：通过 CX8 800Gbps RNIC 接入 InfiniBand RDMA Scale Out 网络，实现多个 NVL72 SuperNode 组成更大规模的 SuperPOD（例如 8 个 DGX GB200 NVL72 组成一个包含 576 块 B200 GPU 的 SuperPOD）。

（二）华为 CloudMatrix 384

CloudMatrix 384 是华为推出的超大规模 AI 超节点解决方案，由 384 颗昇腾 910C NPU 芯片通过全连接拓扑结构互联而成。它创新性地提出了对等计算架构，将总线从服务器内部扩展到整机柜甚至跨机柜。

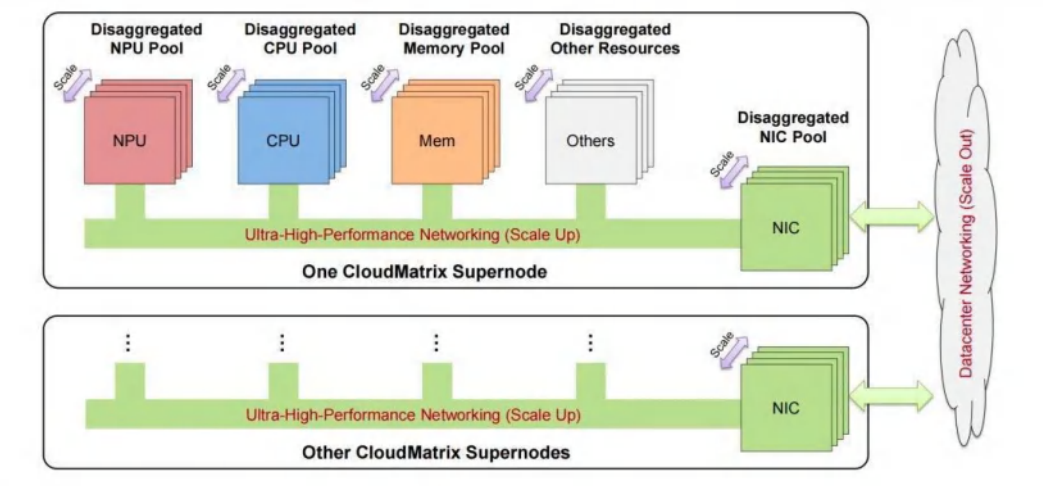

图 13 华为 CloudMatrix 架构

- Compute Tray：每个 Compute Tray 包含 8 块 910C NPU，并且内置了 7 个 L1-HCCS-SW 交换芯片（上联 Scale Up）+1 个 CDR 交换芯片（上联 Scale Out）。910C NPU 采用 Chiplet 技术，集成了 2 颗 910B 和 8 颗 HBM2e 内存，单卡算力达到 FP16 781.25 TFLOPS，内存带宽 3.2TB/s。每张 NPU 基于 HCCS（High-speed Computing Cluster System）GPU-GPU 私有高速互联协议使用 8 个通路分别连接到 L1-HCCS-SW 交换芯片，实现无阻塞带宽收敛比。

- Switch Tray：每个 Switch Tray（CloudEngine 16800 交换机）有 16 个业务槽，每个业务槽最大支持 48 个 400G 接口，整机支持 768 个 400G 接口。采用单层扁平化拓扑，构建 NPU 全互连（All-to-All）拓扑结构，消除传统网络的带宽瓶颈。

- Scale Up：CloudMatrix 384 的 Scale Up 带宽高达 269TB/s，是 NVL72 的 2.1 倍。因物理距离限制，采用 400G 低功耗光模块（LPO），省略了传统 DSP 芯片以降低时延和功耗。

- Scale Out：采用 Spine-Leaf 8 导轨拓扑，通过 400G 光模块构建 Scale Out 网络，实现超节点间的互联，总带宽是 NVIDIA NVL72 的 5.3 倍。

（三）ETH-X

由 ODCC 牵头，联合中国信通院、腾讯等单位发起的 ETH-X 项目可以支持单个超节点 64 卡的计算能力，和英伟达的私有 NVLink 方案不同，ETH-X 采用更为开放的 RoCE 方案。

- Compute Tray：每个 Compute Tray 包含 4 张 GPU 和 1 个 X86 CPU，CPU 和 GPU 之间通过 PCIe Switch 对接。整个机柜共 64 张 GPU。同时每个 Compute Tray 提供 4 个 NIC 用于 Scale Out 方向的扩展。

- Switch Tray：每个 Switch Tray 包含 1 颗支持 RoCE 的高性能 51.2Tbps 以太网交换芯片，整个机柜提供 8 个 Switch 芯片。GPU 和 Switch 芯片支持 100G serdes。

ETH-X 整机柜 GPU 互联带宽为 204.8Tbps。8 个 Switch Tray 支持 409.6Tbps 的带宽，一半用于超节点柜内连接 GPU，另一半的带宽用于背靠背连接旁边机柜的超节点或者通过 L2 HB Switch 做更大的 HBD 域 Scale Up 扩展。Intel Gaudi3 GPU 提供 4.8Tbps 的带宽，整个超节点机柜需要 12 个 Switch Tray。ETH-X 也支持 Switch Tray 没有外部 Scale Up 扩展口的方案，所有 serdes 连接都用于柜内互联，只需要 4 个 2U 高的 Switch Tray。

3.4.3 6G 与 AI 网络协同

作为下一代移动通信技术的核心方向，6G 网络不仅聚焦更高的速率与更广的覆盖，还被寄予原生智能化的期望。因此，6G 与 AI 网络的协同已成为未来网络架构演进的关键议题，两者之间的深度融合，将催生出全新的智能算力网络范式。从应用需求看，AI 网络对 6G 的核心诉求可归纳为：

极致低时延与确定性服务：面向训推场景中跨设备、跨地域的实时数据同步，AI 网络要求网络能够提供亚毫秒级端到端通信延迟，且具备微抖动和高稳定性；

分布式资源感知与协同调度能力：AI 网络的计算资源呈异构、跨域分布，6G 网络需具备原生的资源发现—路径匹配—带宽调配能力，支撑 AI 训练作业或推理服务在边、云间高效调度；

原生多模态感知能力：AI 网络中的数据流不再局限于传统结构化报文，而是包含音频、视频、文本等多种模态，要求底层通信系统具备识别、分类、加速与编解码优化等多模态协同能力。

这些特性构成了 6G 网络演进的目标边界，从通道能力迈向智能算力基础设施的融合平台。反过来看，6G 网络本身的架构演进也将深度嵌入 AI 技术，通过 AI 网络的协同赋能，实现从连接中心向智能中心的范式转变。其技术路径主要体现在以下层面：

网络架构智能化：6G 架构在设计之初就引入 AI 控制与协同模块，网络控制平面使用深度强化学习进行路由、频谱调度与功率控制；数据平面支持边缘推理节点进行本地智能处理；引入智能控制代理，模拟复杂拓扑行为，优化资源编排。

通信计算融合：6G 网络不再将通信与算力视作割裂系统，而是构建通信即计算架构。在边缘节点集成 AI 推理能力，实现边学边用；在链路层引入联合编码/压缩感知机制，减少跨域通信负载。

AI 网络与 6G 协同：AI 网络可感知底层链路负载、拓扑变化，实时调整并行粒度或模型切片方式；6G 网络可根据 AI 作业级 SLA 需求（如延迟、吞吐、能耗）主动执行路径调度、QoS 区分与任务迁移。

6G 与 AI 网络协同并非简单的通信和算法组合，而是面向未来智能社会的核心基础设施重构。二者的融合，将打造真正意义上的智能可编程网络，支撑大模型、数字孪生、类脑计算等未来 AI 应用的高效运行。

第 4 章　Network for AI 典型应用实践

本章精选行业领先的 Network for AI 应用案例，全面覆盖大规模 AI 模型训练、实时推理等关键场景，并延伸至智能制造、智慧教学等垂直领域。通过深度解构这些案例的技术架构，揭示其如何实现网络性能与 AI 需求的精准匹配，提炼可复制、可推广的落地方法论，为行业提供兼具前瞻性与实操性的参考范式。

4.1 移动云新型智算网络架构

移动云新型智算网络架构 HPN1.0 面向超大规模 AI 集群通信场景，采用开放以太网技术路线，通过自主研发的高性能交换设备与 FARE（全自适应路由以太网）协议，突破了传统以太网在集合通信场景下的负载均衡瓶颈，实现带宽利用率和训练效率的双提升，满足智算中心大模型训练与推理场景对高吞吐、低延迟、高可靠的极致要求。

（1）Scale Out 网络

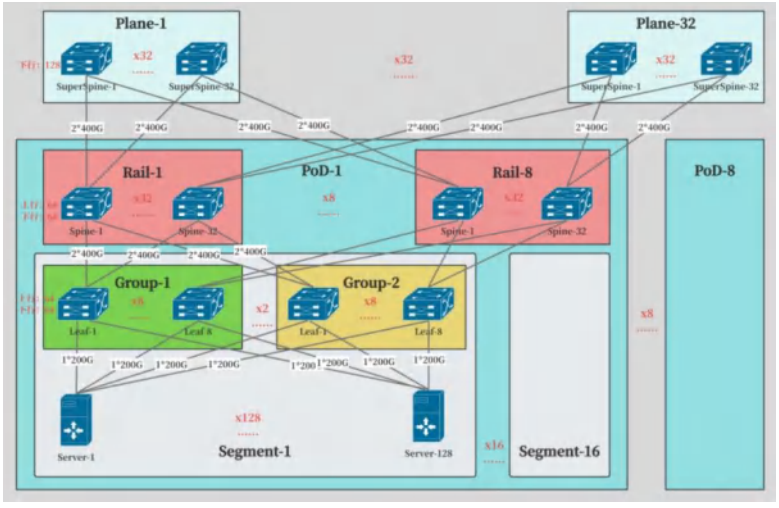

图 14　移动云十万卡集群智算网络架构

- 架构创新：采用三层 CLOS 多轨道组网，单 PoD 支持多达 6 万张 GPU 服务器互联，收敛比达到 15:1，多轨道设计最大化机内网络与机外网络的协同能力，减少跨层通信，降低了整体网络延迟；

- 协议突破：FARE 协议针对 AI 训练中流数少、单流大、高并发的流量特征，支持多路径包喷洒机制，带宽利用率可达 95%以上；

- 极致性能：基于 51.2T 芯片自主研发磐石智算交换机，单台 8 卡 GPU 服务器支持 3.2Tbps 的超高接入带宽；通过流量转发路径的优化、精准流控等手段的综合运用，确保端到端延迟在 10 微秒范围内。

（2）Scale Up 网络

采基于开放以太网技术路线，实现面向超节点的 Scale Up 网络，支持几十卡到上千卡超节点规模，卡间带宽高达 800GB/s，远高于 PCIe Gen5 的 128GB/s。

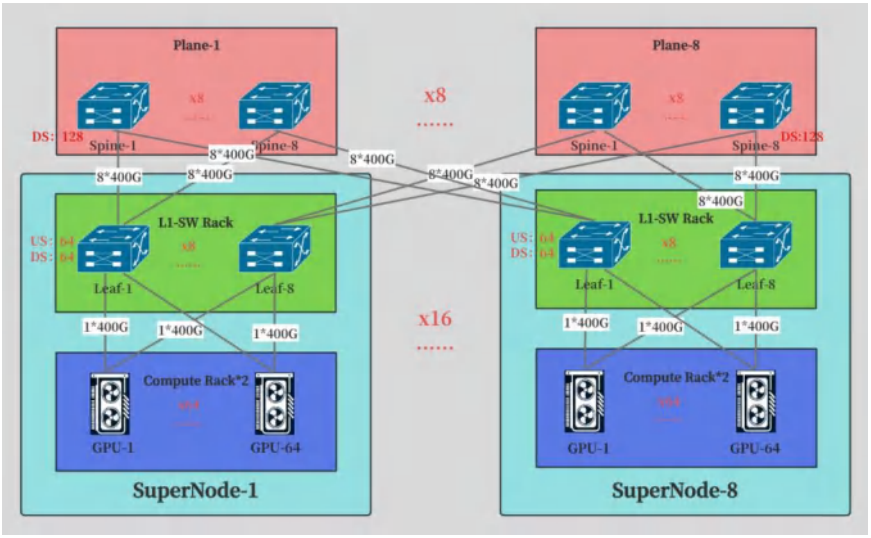

图 15 移动云 1024 卡超节点

- 硬件开放：超节点采用全开放硬件架构，计算与交换节点采用标准化机型，风冷和液冷灵活配置，单机柜功耗在 40～60kW 范围，可以同时适用于训练和推理场景；

- 协议适配：基于优化的 RoCE 协议实现远端内存访问，同时通过适配层支持内存语义访问，兼容 AI 大模型训练通信需求；

- 性能指标：跨 GPU 远端访问延迟控制在 300 纳秒以内。

HPN1.0 通过积木式模块化设计设计思路，采用高度标准化的、跨不同厂商的 GPU 服务器（如 8 卡风冷或液冷 GPU 服务器）与智算交换机（如 51.2T 风冷或液冷智算交换机），通过标准的 AEC 有源铜缆及光纤互联构成。

目前移动云新型智算网络架构 HPN1.0 已在实际智算中心项目中完成落地验证，通过采用标准化 GPU 服务器与开放交换设备，构建了高带宽、低延迟的智算集群网络，具备快速部署、按需扩展、稳定运行等优势。项目实现了千卡规模的超节点部署，训练效率与网络性能表现优异，展现出良好的工程可复制性与多场景适配能力，为后续多地算力基础设施建设提供了可推广的技术范式。

4.2 天翼云智算项目

在生成式 AI 大模型驱动下，教育、医疗、汽车等行业加速应用落地。天翼云顺应 AI 发展趋势，布局大规模 GPU 智算资源池，重点构建高性能 AI 训练集群网络。

（1）核心架构与协议

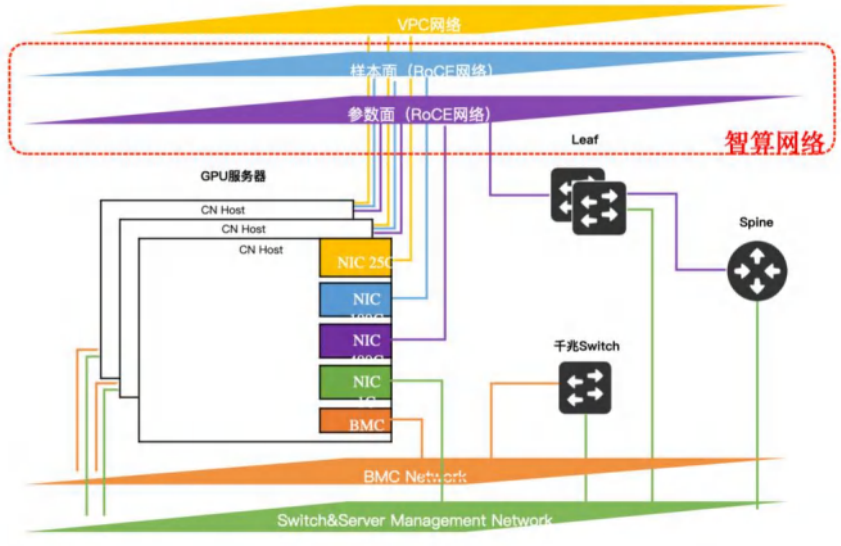

图 16：天翼云智算项目核心架构

项目采用多平面物理组网，聚焦参数面网络实现分布式训练参数同步，构建高带宽、低延迟的 AI 训练集群：

- 底层协议栈：选用 RoCEv2 协议，实现 RDMA 在以太网网络中的传输，仅使用 IB 的“轻量级”传输层，降低设备成本与网络环境需求。
- 无损以太网络：RoCEv2 使用 UDP 头部来封装 RDMA 相关协议栈内容，结合二层 PFC Pause 帧与三层 ECN 标记，确保流量低时延、无损转发。

（2）硬件选型：

- Leaf 层采用华为 4 槽 CE9860 盒式交换机（8 × 400GE 接口）；
- Spine 层选用华为 8 槽 CE16808/16 槽 CE16816 框式交换机（36 × 400G 端口）；
- GPU 服务器搭载昇腾 910B 芯片。

（3）组网设计

参数面和样本面采用二层 Clos 架构，Spine 与 Leaf 采用 Full-Mesh 全互联，运行 eBGP 协议。

参数面：服务器使用 8*200G 接口，单轨接入 Leaf 交换机。Leaf 交换机通过 32 × 200G 端口下行连接服务器，采用 Y 型一分二线缆与服务器 200G 接口对接，共接入 4 台服务器。其中，单台服务器通过 8 个 200G 网口连接至一台 Leaf 交换机，8 个网口分别配置独立的 IP 地址。Leaf 交换机通过 16 × 400G 端口上行连接至 Spine 交换机，Spine 交换机端口扇出决定了 AI 集群规模。例如万卡集群需要至少 313 台 Leaf 接入，则选用 16 台 16 槽的框式交换机，且单台 Spine 设备的 400G 端口数大于 313。

样本面：服务器使用 2*100G 双口接入两台 Leaf 交换机，对于 GPU 服务器的接入，bond 采用 mode 1 主备方式，对于 HPFS 存储服务器，由于可以运行自研操作系统，实现 arp 双发，接入设计上采用去堆叠方案，服务器 bond 模式使用 mode 4。样本面 Leaf 交换机采用 32 × 100G 端口形态，Spine 采用插卡框机，两台 Leaf 设备为一组构成一个 block，一个 block 可接入 16 台服务器。样本面的存储服务器和计算服务器之间按照比例配置。

业务面：智算业务面接入基于天翼云 4.0 架构设计的通算资源池，服务器采用 2*25G 紫金 DPU，并上联接入天翼云自研交换机。

4.3 阿里云 HPN7.0 新型智算网络

阿里云 HPN7.0 是面向 AI 大模型训练场景设计的智算网络架构，其核心目标是通过创新的拓扑设计、多路径冗余和自研通信技术，解决万卡级 GPU 集群的高性能、高稳定性及可扩展性挑战。

（1）架构设计

采用“双上联+多轨+双平面”设计。这种设计能确保网络在超高负载下仍保持高效、稳定运行，满足 AI 大模型对计算资源的高需求。

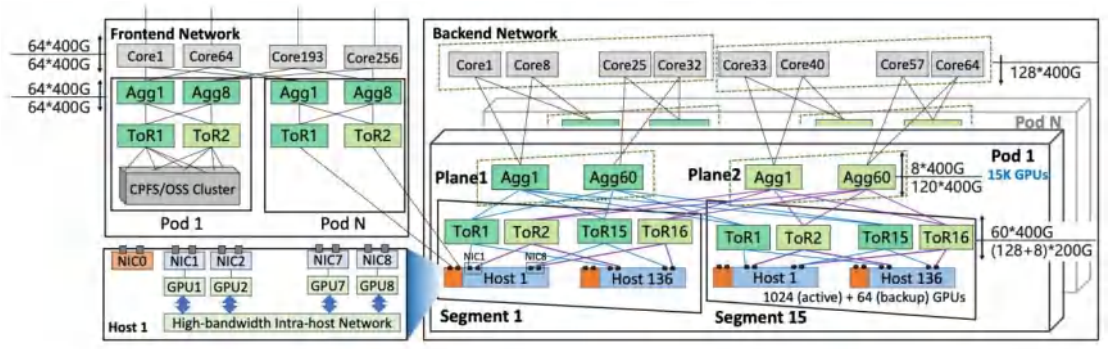

图 17 阿里云 HPN 架构

双上联：每台 GPU 服务器配备双物理网卡（或单网卡双端口），分别连接至不同 Leaf 交换机，形成冗余路径。这种设计提高了网络的可靠性和性能，确

保在任一上联链路或接入层交换机发生故障时，网络流量能够自动切换至另一端口，保障训练任务的连续性和稳定性。

多轨：允许多个数据流并行传输，增加了网络吞吐量。每个 GPU 都与多张高性能网卡相连，通过多轨通信技术实现集群内 GPU 的全互联，优化了长尾时延，为大规模集群计算提供了更加高效和稳定的网络通信支持。

双平面：通过双平面转发机制，将网络流量均匀分配至两个独立的网络平面，降低了哈希极化现象的发生概率，优化了网络流量的分配效率和网络的整体性能。

硬件配置：配备 51.2Tbps 单芯片以太网交换机和 400G 高性能网卡，可实现单层千卡、两层万卡的高性能和高稳定性互联，为 AI 大模型的训练及推理提供强大的硬件支撑。

（2）关键技术

Solar-RDMA 协议：是一种基于 RDMA 的通信协议，能够实现数据的快速传输和高效处理，可在大规模分布式系统中实现低延迟、高带宽的数据传输，确保 AI 大模型的高效运行。Solar-RDMA 还提供了高精度拥塞控制算法，结合网络负载的动态感知，能够实现对数据流级别的精细控制。

ACCL 通信库：是阿里云针对 AI 计算场景专门设计的通信库，优化了 AI 计算过程中的数据传输和同步操作，能显著提升计算的效率和稳定性，为大模型提供稳定可靠的网络通信支持。

自 2023 年 9 月大规模部署以来，HPN7.0 在大模型训练性能方面表现卓越，与上一代架构相比，在典型场景下实现了高达 14.9%的性能飞跃。阿里云通义千问 2.5 版本大模型就是基于 HPN7.0 高性能网络集群训练而成，其中文性能全面赶超 GPT-4 Turbo。

目前阿里云已经推出了下一代训推一体融合网络架构 HPN8.0，旨在支撑万卡到几十万卡的超大规模智算集群。HPN8.0 采用全自研软硬件系统，硬件上包括 102.4T 大芯片交换机、自研 400G/800G/1600G 光模块与硅光芯片等；软件上涵盖 ACCL 通信库（拓展至适配多场景）、Nimitz 容器网络、Stellar-RDMA 协议栈等。架构设计上，Back-end GPU 互联网络通过带宽升级、多平面扩展及协议增强，支持规模扩大 8 倍和跨地域互联；Front-end 网络则实现 N × 10 万级别 GPU 规模覆盖与 AZ 内全互联，对接 VPC 和存储系统。整体凭借超大集群支持、低时延高可靠及全场景适配等优势，进一步突破 AI 智算的网络瓶颈，为超大规模 AI 集群提供核心支撑。

4.4 奇异摩尔 AI Networking 全栈解决方案

奇异摩尔成立于 2021 年初，依托高性能 RDMA、网络控制和 Chiplet 等核心技术，构建基于开放生态的统一互联架构 Kiwi Fabric，为超大规模 AI 智算芯片/平台提供高性能互联产品及解决方案。

奇异摩尔开放统一架构 Kiwi Fabric 的 AI Networking 全栈解决方案，覆盖从数据中心级网间互联、芯片级片间互联到芯片内部的深度互联。其核心优势在于：

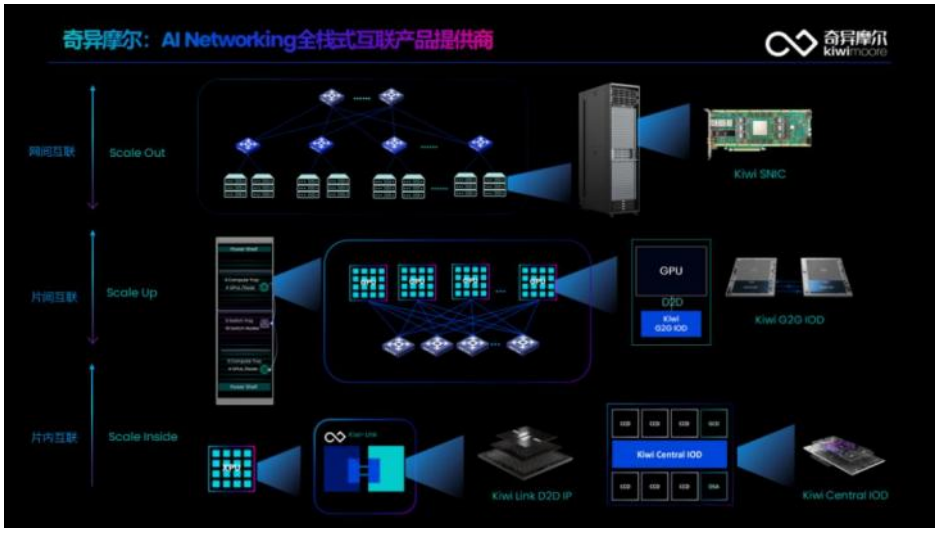

图 18 奇异摩尔 AI Networking 全栈解决方案

（1）Scale Out 网间互联：专为 AI 原生定制的超级网卡

Kiwi SNIC AI 原生超级网卡适用于 AI 大模型训推集群的北向网络互联（网间互联）。产品基于以太网和下一代高性能 RDMA 技术，内建高性能 RDMA 数据传输引擎和自适应网络调度算法，可实现 Tb 级高速互联、和十万卡级网络拓扑。

（2）Scale Up 片间互联：构筑 AI 网络超节点的互联芯粒

Kiwi G2G IO Die 超节点互联芯粒是国内少有的开源&通用化超节点互联方案；该产品为 Chiplet 形态，通过先进封装集成在 XPU 计算芯片内，通过网络接口和交换机互联，支持 1K+ AI 网络超节点，实现 xPU 芯片间的 TB 级超高速互联。芯粒内建高性能数据传输引擎和可编程网络控制引擎，支持内存和消息双语义，多种超节点协议，多种拓扑结构。

（3）Scale-Inside 芯片内互联：破局算力瓶颈，打造高性能芯片

Kiwi Central IOD & 3D base Die 属于行业前沿的 Chiplet 芯粒产品；分别对标 AMD Zen 系列 CPU IO 芯粒以及 Intel Metor Lake 异构芯片内 Base Die（基于 3D 先进封装）。

4.5 第一线助力教育企业私域 AI 落地方案

某教育企业拥有超 300TB 的多元数据，亟需构建私域空间实现数据存储、模型训练与部署。第一线 DYXnet 为其打造私有向量数据库知识库，结合隐私计算算力完成模型微调，确保数据全流程本地化处理。同时，通过打通企业原有存储与边缘节点，构建安全互联网络，调度 64 台算力资源完成模型训练，并快速切换至分布式推理资源，依托 100Gbps 高速 AI 内网，实现数据远程训练与模型分布式部署，最终将新模型部署至公网服务用户。第一线通过 AI 原生超互联总线架构，为方案提供技术支撑：

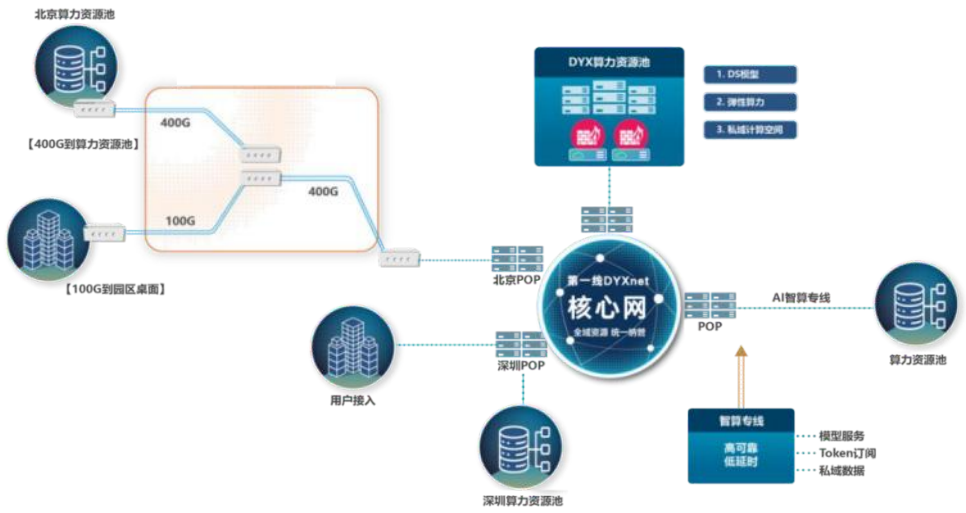

图 19 第一线 AI 原生超互联总线架构

- 网络接入与核心链路：接入侧支持多终端差异化接入，如家庭/企业通过全光直连/F5G、移动终端通过 5G 网络切片接入；核心层与运营商协作打通关键链路，实现高速安全组网；管理层将园区网络控制面向上迁移，结合区块链技术支持企业整体私域网络管理，简化运维。

- 底层安全隔离：采用 VXLAN、FlexE、SRv6 等技术，基于用户维度隔离网络，结合数字身份与密码学，保障数据传输与存储安全。

- 广域网优化：融合生态伙伴方案，运用加密流识别与智能调度技术，精准处理智算业务，实现千万级流量均衡调度，提升网络吞吐与传输可靠性。

- 远程 RDMA 创新：在架构中应用远程 RDMA 协议，大幅提升数据传输与模型训练效率，如 2000 公里以上传输速度达 TCP 的 20 倍，10km 节点协同训练差距控制在 2%以内。

- 无盘隐私计算：计算服务器采用无盘设计，数据算后即清，通过动态私有连接与内存计算，确保数据仅存在于当前计算空间，保障隐私安全。

4.6 微众银行金融级智算 AI 网络建设与实践方案

在人工智能驱动行业变革的浪潮中，微众银行率先提出向“AI 原生银行”转型的战略目标，构建了覆盖 AI 基础设施、应用与治理的三层能力体系。面对大模型时代算力网络的核心挑战，微众银行于 2025 年推出金融业首款自研交换机 WB3000，打造“白盒硬件+开源系统+自研智能管控”全栈自主可控的智算网络解决方案，为千卡级 AI 训练与推理提供高速网络底座，助力金融服务迈向智能化新阶段。微众银行以分层解耦架构破局，实现性能与自主可控的双重突破：

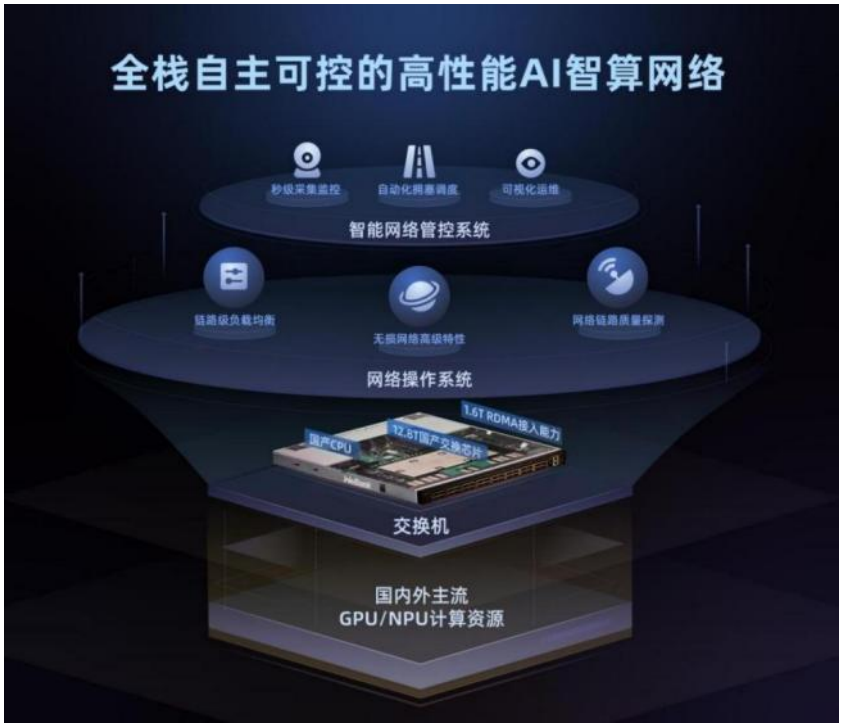

图 20 微众银行 AI 智算网络

- 硬件层革新：基于国产 12.8T 交换芯片与信创 CPU 打造白盒交换机 WB3000，支持 32 个 400G 端口和 1.6T RDMA 接入能力，适配主流 GPU/NPU 算力卡。通过核心部件 100%国产化与非核心部件替代清单，建网成本降低 70%，打破商用方案垄断；

- 系统层创新：基于开源 SONiC 深度定制 WeNOS 网络操作系统，首创链路级负载分担组件 Link-SLB。通过预配置确定性哈希路径解决 ECMP 拥塞问题，

实测集合通信带宽提升 40%，并实现毫秒级故障切换。微众银行由此成为全球首家入选 SONiC 贡献组织的金融机构；

- 智能管控突破：自研管控系统融合 Telemetry 秒级采集与 sFlow 流量分析，实时感知端口拥塞并自动调度至最优路径。结合 AI 训练任务动态回收策略，构建“采集-定位-调度-回收”全闭环智能运维体系，大幅降低人工干预需求。

4.7 益思芯创新智能网卡解决方案

益思芯科技精准切入制约 AI 集群效率的诸多痛点，打造了全系列基于 FPGA 的智能网卡产品，为 AI 场景提供强大网络加速引擎。其技术核心在于 P4 可编程网络处理能力与深度硬件卸载架构：

- 创新的自主知识产权 DSA P4 引擎支持灵活定制 AI 工作流，实现网络功能的动态优化；

- 全硬件加速的 RoCEv2 引擎为 AI/HPC 集群提供超低延迟 RDMA 网络，突破跨节点通信性能限制；

- 云原生 NVMe-oF 共享存储加速引擎则通过硬件卸载 NVMe 协议，将本地存储的高性能扩展至网络共享环境，大幅提升分布式 AI 训练的数据吞吐效率。

益思芯产品如 Stargate-R2100 RDMA 智能网卡（2x100G）及 Stargate-S1100 存储加速卡，已在云厂商和智算中心实际部署，验证了其在 AI 模型训练与大数据分析场景中显著降低时延、提升带宽利用率的实效。该方案的优势不仅在于单卡性能，其完整的智能网卡产品线覆盖从 25G 到 100G，全面对标国际领先水平，更构建了面向 AI 的开放生态，无缝集成 DPDK/SPDK 开源库，支持云原生驱动及国密安全算法，为 AI 基础设施提供从网络通信、存储访问到数据安全的端到端加速。

第 5 章　Network for AI 未来发展及展望

本章前瞻性探讨 Network for AI 未来发展趋势，为产业发展提供战略性思考与展望。

5.1　未来发展趋势

（一）网络与计算深度融合

网络与计算的深度融合将成为推动 AI 发展的关键力量。随着 AI 应用的不断拓展，数据传输与处理的实时性需求愈发迫切，促使网络从单纯的数据载体向具备强大计算能力的智能平台转变。网络边缘将具备实时处理海量数据的能力，减少数据传输延迟，推动自动驾驶、工业互联网等时延敏感的应用走向成熟。计算与网络的深度融合，将实现算力资源的动态调配与高效利用，用户通过网络即可便捷获取所需算力，加速 AI 应用的落地与普及。

（二）AI 网络定制化演进

网络将针对 AI 应用的多样化需求，构建更具适配性的支撑体系。针对 AI 训练阶段对海量数据传输的需求，网络将通过优化传输协议、提升带宽利用率，实现大规模数据集的高效分发与同步，缩短模型训练周期。对于 AI 推理场景，网络会重点保障低时延与高可靠性，通过动态调整路由策略、优先调度推理请求等方式，确保实时交互类 AI 应用的流畅运行。同时，网络将建立弹性伸缩机制，根据 AI 任务的算力需求变化，自动调整资源分配，为 AI 模型的迭代优化和规模化部署提供稳定、高效的底层支撑。

（三）生态开放与协同创新

生态的开放与协同创新是未来发展的核心驱动力。硬件层面，各厂商联合研发适配 AI 计算与网络传输的高性能硬件，提升系统整体性能与效率。软件方面，

开源社区汇聚全球开发者智慧，加速 AI 框架、算法与网络协议的创新迭代，推动技术的快速发展与广泛应用。跨行业的深度合作将催生更多创新应用，利用 AI 网络提升服务质量与效率，促进社会各领域的数字化、智能化转型。

（四）绿色低碳网络基础设施

绿色低碳的网络基础设施将成为未来发展的重要方向。主要路径包括：架构级节能，通过去中心化算力调度、液冷/自然冷却技术、高集成度设备降低基础能耗；全生命周期绿色化，从芯片设计、设备制造、到数据中心规划贯彻绿色理念；同时，加大对太阳能、风能等可再生能源的利用，为网络基础设施提供可持续的能源支持，助力实现网络发展与环境保护的双赢。

5.2 未来展望及建议

（一）对技术研发的建议

（1）推动网络赋能 AI 的技术标准化与模块化：制定网络支撑 AI 应用的通用技术标准，明确实时推理、大规模训练、分布式协同等不同 AI 场景下的网络性能量化指标，统一接口协议与数据交互格式。同时，研发模块化的网络功能组件，如可插拔的 AI 任务适配模块、标准化的资源调度插件等，让不同行业的 AI 开发者能快速调用网络功能，无需关注底层技术细节。通过标准化与模块化降低技术适配成本，加速网络赋能 AI 的规模化落地。

（2）突破关键技术瓶颈：针对 AI 网络中的算力、存储、传输等关键环节，持续投入研发，突破性能瓶颈。例如，研发面向 AI 计算的专用网络芯片和设备，提升网络传输带宽和降低时延；探索新型网络架构，如光电融合网络、量子网络等，为 AI 应用提供更强大的基础设施支撑。此外，还需关注 AI 模型的小型化、轻量化技术，使其能够在边缘设备上高效运行。

(3) 注重跨领域技术协同创新：打破网络技术与 AI 研发、垂直行业应用之间的壁垒，搭建集技术研发、测试验证、成果转化于一体的协同平台。推动网络通信专家、人工智能专家、行业应用开发者共同参与，联合定义网络需求指标，协作开发适配性技术方案。例如，在工业 AI 领域，网络团队与制造业专家合作，根据生产线的实时性要求定制低时延网络协议，同时结合 AI 算法优化设备数据的采集与传输策略，实现技术创新与行业需求的精准对接。

(二) 对产业发展的建议

(1) 构建开放共赢的产业生态：鼓励产业链上下游企业加强合作，共同打造开放、协同、共赢的 AI 网络产业生态系统。支持开源社区发展，推动 AI 网络相关技术标准和规范的制定，降低技术门槛，促进技术普及和应用。通过建立产业联盟、合作平台等形式，汇聚各方力量，共同推动产业创新和发展。

(2) 拓展多元化应用场景：推动 AI 技术与实体经济深度融合。例如，在智能制造、智慧城市、自动驾驶、医疗健康等领域，结合具体业务需求，开发定制化的 AI 网络解决方案，创造新的商业模式和增长点。通过示范项目和标杆案例的推广，加速 AI 网络技术的普及和应用。

(3) 加强国际交流与合作：积极参与全球 AI 网络技术标准和产业发展规则的制定，提升我国在国际 AI 网络领域的影响力。鼓励国内企业与国际领先企业、研究机构开展技术交流和项目合作，共同应对全球性挑战，实现互利共赢。

第二部分 AI for Network：AI 赋能的网络智能化升级

第 6 章 AI 驱动的网络智能化发展

在数字化转型纵深推进的背景下，网络作为连接物理世界与数字世界的核心枢纽，正面临规模与复杂度增长、业务需求多元化、安全威胁复杂化等多重挑战。传统依赖人工配置的静态网络架构与被动响应式运维模式已难以支撑动态异构的网络环境运行需求。这一矛盾驱使网络技术范式发生本质跃迁，从以连通性为核心的连接型网络，向以认知自治为目标的智能型网络跨越。本章将系统剖析网络智能化发展的内在驱动力和升级流程，为全面理解 AI 赋能网络的深层逻辑与未来愿景奠定理论框架。

6.1 网络管理的挑战

（一）网络规模与复杂度持续增长

随着 5G/6G、物联网、大数据等技术的快速发展和广泛应用，全球网络规模正呈现爆发式增长。网络所承载的业务类型、服务对象、接入设备均向多元化演进，这种规模和复杂度的激增导致网络拓扑日趋繁杂、协议种类愈发多样、人工管理配置难度呈指数级增长。因此，如何有效应对日益增长的网络复杂性，已成为推动网络智能化升级的首要驱动力。

（二）运维效率与成本控制压力

网络规模的持续扩大与架构的日益复杂，使得传统运维陷入效率与成本的双重困境。传统以人工为主、依赖静态规则和专家经验的运维模式，响应速度慢、故障定位时间长且需要投入大量专业人力，在面对海量告警、复杂故障和动态业

务需求时，已显得力不从心。运维效率滞后、成本高昂与服务质量不稳定的矛盾愈发突出。为突破这一困境，网络运维亟需从事后救火的被动响应模式转向事前预警的主动防御体系，从人工干预转向自动化智能化范式，以实现网络资源的优化配置和高效利用。

（三）业务体验与网络性能优化需求

在数字化时代，用户对网络服务的期望越来越高，对业务体验的感知也越来越敏感。在线游戏、自动驾驶、工业互联网等业务对网络带宽、时延、抖动和可靠性提出了极致要求。传统网络在面对突发流量、局部拥塞或设备故障时，往往难以快速响应并进行自适应调整，从而导致业务体验下降、卡顿甚至中断。智能化网络能够实时感知业务流量特征，精准识别应用类型，并通过智能流量调度和动态路径规划等技术，将资源按需实时地分配给最需要的业务，从而确保关键业务的极致体验。

（四）安全威胁防护的智能化需求

网络边界的模糊化和攻击手段的多样化，使得网络安全防护面临严峻挑战。高级持续性威胁、零日漏洞、分布式拒绝服务攻击等新型网络攻击具有隐蔽性强、传播速度快、破坏力大等特点。为了有效应对不断演变的安全威胁，网络安全防护需要从静态防御向动态感知和自动化响应转变。这种智能化的安全能力可以更快地发现未知威胁，自动执行隔离和阻断策略，构建起一个能够自我学习、自我适应的动态安全防护体系。

6.2 网络智能化演进体系

在网络智能化演进浪潮中，业界普遍采纳了自智网络等级划分的评估体系，等级越高代表网络的自动化和智能化程度越强，这一标准为网络智能化的发展提供了清晰的演进路径和目标。

自智网络等级	L0: 人工运维	L1: 辅助运维	L2: 部分自智网络	L3: 条件自智网络	L4: 高度自智网络	L5: 完全自智网络
执行	P	P/S	S	S	S	S
感知	P	P/S	P/S	S	S	S
分析	P	P	P/S	P/S	S	S
决策	P	P	P	P/S	S	S
意图/体验	P	P	P	P	P/S	S
适用性	N/A	选择场景				所有场景
P 人(手工) S 系统(自主)						

图 21 自智网络等级划分

- L0 人工运维：这是网络运维的最初阶段，系统仅能提供基础的辅助监控功能，所有关键操作，无论是配置下发还是状态查询，都完全依赖人工通过命令行完成。网络管理效率极低，且极易因人为操作失误引发各类问题，难以适应网络规模扩大的基本需求。

- L1 辅助运维：针对网络运维中具有明确规则、重复性高的任务，通过专门的工具或脚本实现批量操作。这一阶段借助工具，在一定程度上减轻了人工负担，提高了网络运维管理工作的执行效率和用户对网络管理的感知效率，但本质上仍依赖人工定义的规则，智能化程度有限。

- L2 部分自智网络：系统能够依据人工预先定义的策略，辅助用户实现部分网络运营管理工作流程的闭环操作，最终的决策权力仍掌握在用户手中。网络在部分场景下展现出一定的自主性，但整体仍受限于人工设定的框架。

- L3 条件自智网络：系统的智能化分析能力得到显著提升，可以自动感知网络状态与资源信息，还具备事前评估、事后自动验收以及问题自动定位等能力。基于人工定义的闭环自动化策略，系统能够实现部分特定场景的闭环管理，网络自主性进一步增强。

- L4 高度自智网络：相较于 L3 阶段，其网络智能化程度实现了跨越式提升。系统能够主动感知网元及整个网络的状态，通过趋势分析预判潜在风险，并主动采取优化措施，确保网络持续满足业务需求。

- L5 完全自智网络：这是自智网络发展的终极目标。此时的网络能够实现完全自主运行，无需任何人工干预，系统具备完全自主的决策能力，可实现自我演进、自我适应、自我修复和自我优化。

当前，自智网络产业正处于从 L3 条件自智网络加速迈向崭新的 L4 高度自智网络的阶段。根据 TM Forum 的评估，全球 91%的运营商已制定自智网络长期战略并追加投资，70%的通信服务提供商将投资网络基础设施，以实现自动化。《自智网络产业白皮书 6.0》中明确提出面向 2030 年分两个阶段实现 L4 目标：2025 年到 2027 年实现单域“维优营”场景自智闭环，2028 年到 2030 年实现跨域复杂场景端到端闭环。中国三大运营商更是进一步聚焦于高度自智网络的顶层设计，计划在 2025 年初步实现高价值场景的 L4 高度自智目标。L4 自智网络的核心特征主要体现在以下几个方面：

- 意图驱动：网络能够精准理解用户或特定业务的高阶意图，并将这些意图自动、智能地分解转化为网络可执行的低阶策略和指令，整个过程无需人工干预翻译，即可驱动业务目标的实现。

- 全生命周期闭环控制能力：系统能够对网络状态进行实时监测，结合智能根因分析做出自主决策并执行，形成涵盖规划、部署、运维、优化等网络全生命周期的端到端闭环控制，大幅提升网络管理的效率和精准度。

- 跨域协同与优化能力：L4 自智网络能够有效协同和优化不同网络域的资源和功能，打破了传统网络管理的壁垒，实现了网络全局最优配置。

- 预测性维护与自愈能力：借助大数据分析和 AI 模型，系统具备强大的预测能力，能够预判潜在的网络故障和性能劣化趋势，并主动采取预防措施或自主进行修复。

6.3 网络智能化升级流程

网络的智能化升级是一个以 AI 技术为核心驱动力，通过“全域感知、智能分析、自主决策、执行与保障”四个阶段形成完整闭环的动态演进过程。这一流程从对网络全域状态的精准感知入手，依托 AI 智能算法对海量数据进行深度分析与挖掘，进而基于分析结果生成最优化的网络控制策略与资源调度方案，随后精准、自动化地实施决策指令，动态调整网络行为，同时辅以全流程的隐私保护，环环相扣、持续迭代，最终实现网络的智能化跨越。

6.3.1 全域感知

全域感知是网络智能化升级的基石，通过对网络中的设备状态、网络流量、环境状态等进行全方位、实时的数据采集和状态监测，为后续的分析、决策、执行提供数据支撑。

网络流量是网络运行状态的直接反映。传统流量分析依赖端口、协议等静态特征，难以应对复杂网络环境和加密流量挑战。AI 的引入，特别是机器学习和深度学习算法的应用，使得网络能够从海量的流量数据中自动学习和提取高维特征，从而实现对流量模式的智能识别、异常行为的精准检测以及应用类型的细粒度分类。

时空数据进一步丰富了感知的维度。网络是一个不断演化的复杂系统，其状态不仅随空间变化，也随时间演进。时空序列预测模型和时空图神经网络(STGNN, Spatio-Temporal Graph Neural Network)等技术能够捕捉数据随时间、空间演进的依赖关系，通过对历史数据的学习，精准识别正常波动与异常模式，揭示网络行为的内在规律，预测潜在风险。

在网络持续运行的过程中，会实时生成海量且异构的数据类型，这些数据呈现出非结构化、结构化、时序化、图形化等不同模态特征，单一模态数据的孤立分析往往只能捕捉局部信息，难以全面反映网络的复杂状态。基于深度神经网络

（DNN，Deep Neural Networks）的多模态数据融合方法在多个领域展现出优异的性能。多模态数据融合基于深度学习的四维分类体系，提出编码器—解码器模型、注意力机制、图神经网络与生成神经网络四类数据融合方法。

- 基于编码器—解码器模型的方法中，编码器将输入数据转化为保留关键语义的潜在表征并过滤噪声，解码器再据此生成预测结果。

- 基于注意力机制的方法通过为输入数据各部分分配差异化权重，精准提取任务相关信息，在不增加计算成本的前提下提升预测精度。

- 基于图神经网络的方法则在图结构构建阶段直接融合多模态数据，区别于先提取特征后融合的模式，将融合前置到表征学习前。

- 基于生成神经网络的方法借助生成模型的表征能力，将不同模态映射到共享潜在空间，通过对抗训练或联合优化实现模态特征的对齐与互补。

6.3.2 智能分析

智能分析通过数字孪生、实时仿真与预测推演以及因果推理与根因定位等技术对感知层获取的数据进行深度挖掘与认知转化，从而揭示网络运行的深层规律、诊断复杂故障、预测潜在风险，并为后续的智能决策提供科学依据。

数字孪生通过构建物理网络的全要素虚拟映射，将感知层采集的数据统一注入孪生体，形成“物理网络-虚拟孪生体”的实时同步机制。孪生体能够实时、高保真地映射物理网络的运行状态、性能指标、故障情况等，并通过仿真模拟和数据分析，实现对物理网络的实时监控、仿真和预测，为网络优化、故障排查、安全防护等提供更全面的决策分析。

基于数字孪生构建的虚拟空间，实时仿真与预测推演通过动态推演能力，使用户能够很好地模拟、选择、优化解决方案，最终将它们部署到实际网络中，这将降低对实际网络的影响力，减少一定的安全风险。同时可对网络未来状态进行

多维度模拟，利用大数据处理和建模技术实现对现状的评估、对过去的诊断和对未来的预测，模拟各种可能性，以提供更全面的决策分析。

因果推理与根因定位是网络智能运维的核心能力之一。在复杂且动态变化的网络环境中，故障或性能问题往往不是孤立发生的，而是由多个因素相互作用、层层递进导致的。传统基于规则或相关性的分析方法在面对海量告警和复杂故障时，往往难以准确识别问题的根本原因，导致故障排除效率低下，甚至误判。因果推理技术通过挖掘事件间的因果机制而非统计相关性，为网络故障诊断提供了从“现象描述”到“本质溯源”的认知升级，而根因定位则依托因果关系链实现故障源头的精准追溯与预防性治理。通过 AI 技术，网络能够从海量数据中洞察事件的深层因果关系，实现从发现问题到解决问题的智能闭环，从而显著提升网络的韧性、稳定性和运维效率。

6.3.3 自主决策

决策的本质是将分析环节产生的洞察转化为跨域、动态、最优的网络操作策略。通过与知识图谱、意图驱动网络、大模型等技术融合，决策已逐步由依赖专家经验转向数据驱动、自主演化的智能体范式，朝着可泛化、可适应、可解释的方向演进。

知识图谱是一种以图结构描述网络实体及其关系的语义网络模型，通过图结构精准映射网络实体（如路由器、基站、终端设备）与实体间关系，将网络运维、资源调度等场景所需的碎片化信息转化为结构化知识网络，为 AI 驱动的网络决策提供具象化知识支撑。在技术实现上，知识图谱构建主要包括知识抽取、知识融合、知识存储和知识应用四个环节。

意图驱动网络旨在通过主动理解用户意图，将用户或业务的高层级目标转化为可执行的网络配置策略。传统网络决策中，决策目标往往局限于网络自身的技

术指标，与业务意图的衔接需要人工介入，容易导致决策与实际需求脱节。而意图驱动网络通过自然语言处理、知识图谱等手段，直接解析业务方的抽象意图，并将其转化为网络可理解的量化指标。这些指标会作为协同决策的核心约束条件，贯穿多域协同、动态适配的全过程，确保决策结果从源头就与业务目标保持一致。

尽管 AI 在决策方面展现出强大能力，但在当前阶段，完全的自主决策仍面临挑战，尤其是在涉及高风险、高价值或需要人类经验判断的场景。因此，人机共生决策是一个重要的研究方向。AI 可以作为人类专家的强大辅助，提供多维度的分析结果、推荐决策方案，甚至模拟不同决策方案可能带来的影响，从而帮助专家做出更明智的判断。同时，人类专家的反馈和修正也可作为 AI 模型持续学习和优化的重要输入，形成一个正向的决策闭环。这种人机协同模式，既发挥了 AI 在数据处理和模式识别方面的优势，又保留了人类在复杂情境判断和伦理考量等方面不可替代的作用。

6.3.4 执行与保障

执行与保障环节是将协同决策转化为实际网络动作，并确保整个智能化体系稳定运转的“最后一公里”。执行与保障旨在构建一个高度自动化、可编程、可验证的体系，确保网络能够快速、精准地响应智能决策，并持续维持其健康高效的运行状态。

执行的核心在于将生成的决策或策略通过自动化手段精准地部署到网络设备和系统中，这要求网络具备高度的可编程性和自动化能力。SDN 和 NFV 为网络的自动化执行提供了坚实的基础。SDN 通过将控制平面与数据平面分离，实现了网络的集中控制和可编程性，使得 AI 决策可以直接通过控制器下发到网络设备。NFV 则将网络功能从专用硬件中解耦，以软件形式运行在通用服务器上，这使得网络功能的部署、伸缩和调整变得更加灵活。基于此，可以利用自动化工

具或自定义的编排引擎，将复杂的网络配置、策略调整、服务部署等操作封装成可重复执行的自动化流程。

保障是确保网络在智能执行后能够持续稳定、高效运行的关键环节，它涵盖了对执行效果的验证、网络状态的持续监控以及异常情况下的自我修复。意图执行保障对齐用户意图与系统行为，通过策略一致性验证和效果实时反馈，保障精准执行的可靠性。策略一致性验证确认系统中各个组件或代理的策略之间不存在冲突，并且与系统的整体目标保持一致。执行效果实时反馈通过构建“指标感知-偏差识别-根因分析-动态调整”的实时响应链路，及时纠正偏差并优化后续决策。

隐私保护是网络智能化升级全流程面临的一大挑战，如何在利用多方数据的同时，有效保护数据隐私和安全，是一个亟待解决的问题。以联邦学习为代表的隐私保护联合分析技术，为这一难题提供了切实可行的解决方案。作为一种分布式机器学习框架，联邦学习的核心优势在于让各参与方无需共享原始数据的前提下，通过加密的模型参数或梯度交换联合训练全局模型。在实现数据“可用不可见”的同时，充分释放多源数据的联合价值，为网络智能化的合规演进提供了技术支撑。

第 7 章　AI 赋能网络的关键技术

网络智能化要求网络具备自主感知、决策、执行和进化能力，无需人工干预即可适配复杂动态的业务需求与网络环境。意图驱动网络、数字孪生网络以及智能网络大模型被视为网络升级的重要技术抓手。这些技术通过引入智能化、自动化和数据驱动的手段，推动网络向更高阶段演进。

7.1　意图驱动网络

7.1.1　意图驱动网络的定义和架构

（一）意图驱动网络的定义

意图驱动网络是一种以用户意图为核心驱动的网络管理模型，旨在通过自动化、智能化手段实现网络的配置、优化与维护，其核心思想是将用户的业务目标或需求抽象为高层次的策略，并通过算法和系统自动将这些策略转化为具体的网络操作指令，确保网络状态与用户意图的一致性，最终提升网络管理效率，降低运维复杂度。

意图驱动网络的核心理念体现在以下方面：

- 以用户意图为核心：网络的设计与运行以用户的业务需求为导向，通过分析用户意图，自动调整网络资源分配和策略。

- 自动化与智能化：利用人工智能和大数据等技术，实现策略的自动生成、优化和执行，减少人工干预。

- 全场景网络自治：基于网元、资源、服务与业务四个管理层次，构建分层协同的体系化能力，支持跨域网络的统一管理和动态调整，从而提升网络的灵活性与适应性。

（二）意图驱动网络的体系架构

意图驱动网络架构自顶向下主要分为三层，分别业务应用层、意图使能层和基础设施层。各层通过标准化接口实现交互，形成业务意图驱动，网络能力响应的协同体系。

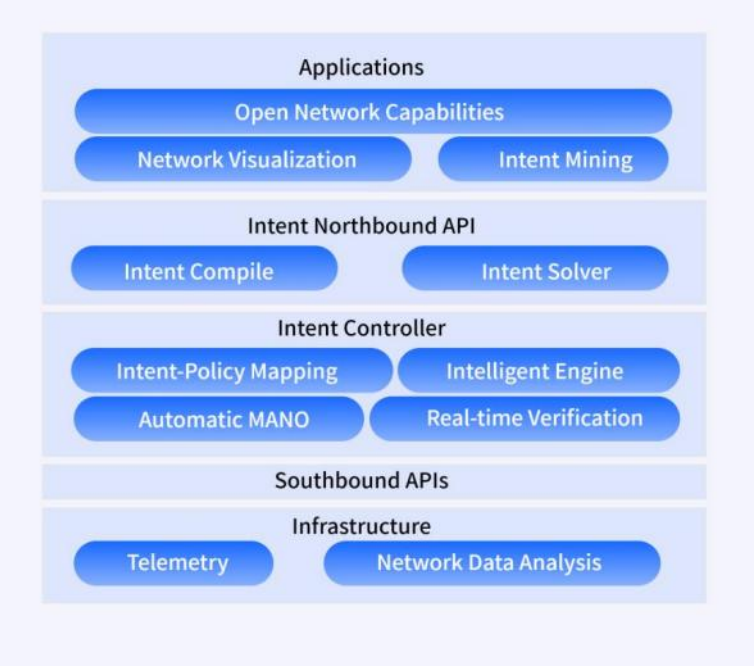

图 22 意图驱动网络架构

业务应用层主要负责生成意图，覆盖不同场景下的各类服务需求。意图可分为直接和间接两种：直接意图是面向管理平面的网络管理需求，可通过应用层直接表达；间接意图则强调用户个体意图，这类意图通常隐藏在用户操作或业务系统运行中，需通过分析行为数据自动提取。业务应用层通过意图使能层提供的编程接口对底层设备进行编程，抽象化网络元素，并通过管理接口实现业务创新多样性。

意图使能层是意图驱动网络的核心，具备管理控制与策略决策能力，主要包括意图策略映射、管理与编排系统、智能引擎和闭环验证四部分。意图使能层接收来自北向接口的意图流，将其转化为当前网络可执行的规范化意图请求后，通过算法将意图中的抽象需求映射到具体的网络资源。基于意图的管理与编排系统可实现资源的统一调度，并通过闭环管理实现网络设备全生命周期监控。借助智

能引擎完成数据收集、数据存储、数据处理、模型训练和参数调整等功能，为策略制定提供先验经验。同时，闭环验证确保了输出的网络配置参数的可靠性。

基础设施层包含各类物理设备实体，并部署大量网络数据采集工具，负责提供反馈信息和策略执行所需的基础资源。

北向接口是连接业务应用层和意图使能层的意图转换模块。意图编译和意图求解器实现意图的表示和一致性检查。南向接口基于虚拟化技术连接各类网络元素设备，主要用于基础设施层和意图使能层之间的交互，并虚拟化计算和通信资源，实现灵活分配。

7.1.2 意图驱动网络的关键技术

意图驱动网络的实现是一个从用户需求输入到目标动态达成的闭环过程，首先需要获取到用户所提出的网络需求（即意图），将接收到的意图转译成网络策略，并根据当前的网络状态验证策略的可执行性，之后将通过验证的策略下发到实际网络中。此外，系统还要实时地监控网络状态，确保用户意图正确实现，并将结果反馈给用户。

（一）意图获取

意图获取的核心任务是准确捕获用户或系统的业务目标与需求，并将其转化为网络系统可处理的初始输入。由于用户意图的表达形式多样，可能涉及不同层级的业务目标或存在语义模糊等问题，因此意图获取需要兼顾灵活性、准确性和全面性，确保网络真正理解用户想要什么。

从意图的来源来看，意图获取覆盖了“外部输入”与“内生生成”两大渠道。外部输入意图主要来自用户的主动表达，是最常见的意图形式。用户可通过自然语言、图形化界面、API 接口等多种方式提交需求，意图获取模块需要兼容这些多样化的交互方式，并从中提取核心需求。内生生成意图则是网络系统基于自身

感知与分析自动产生的目标，体现了网络的自主性。这类意图通常源于对网络状态、业务运行数据的监测。内生意图的获取依赖于网络的全域感知能力，通过收集设备状态、流量变化、故障告警等实时数据，结合历史趋势分析，主动识别潜在需求，从而实现未雨绸缪的智能管理。

意图获取的质量直接影响整个意图驱动网络的效能。若意图捕获不完整，可能导致策略生成偏差；若语义理解错误，则可能引发网络操作失误。因此，这一环节通常会结合机器学习模型持续优化，通过分析历史意图的执行效果，将反馈信息用于感知与预处理算法的迭代升级，逐步提升意图获取的准确性与鲁棒性。

（二）意图分析与转译

意图分析与转译负责将捕捉到的高级、抽象的意图转化为底层网络可执行的具体策略，是一个由抽象向具体转化的过程。在意图分析阶段，核心目标是深入理解意图的本质需求，为后续转译提供清晰的目标蓝图。系统会对标准化意图模型中的字段进行拆解，提取关键要素，包括目标对象、期望指标、约束条件以及优先级。意图分析还需处理潜在的冲突与依赖关系。当多个意图同时存在时，系统需通过冲突检测机制识别矛盾点。

在完成意图解析后，将进入“从目标到操作”的转译过程，即根据分析结果生成网络可执行的细粒度策略。这一过程需结合网络拓扑、资源状态、业务特征等实时数据，通过算法模型将抽象意图映射为具体配置。策略有不同程度的抽象，越往上层抽象程度越高。从 SNMP 到 PBNM 再到 IDNM，策略抽象程度的递进本质上是技术屏蔽能力的升级，逐步剥离底层设备的型号差异、协议细节、拓扑结构等繁杂信息，最终让用户只需关注“需要网络提供什么服务”，而非“网络如何提供服务”。

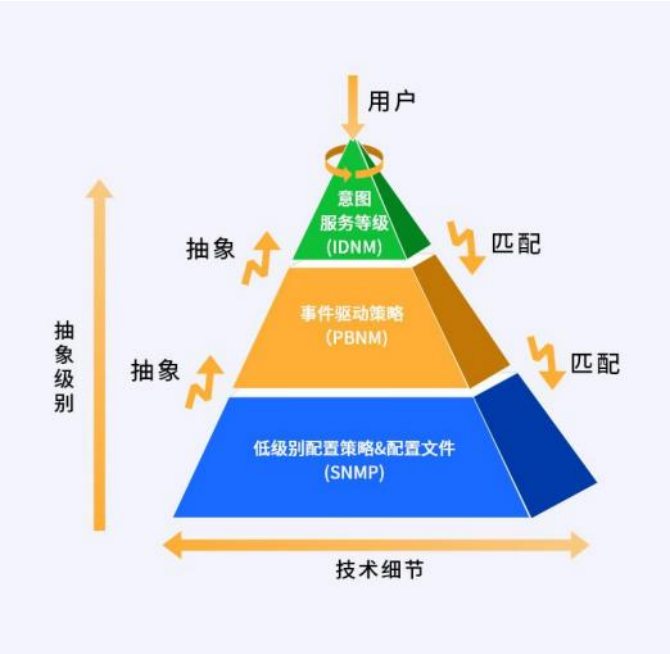

图 23 意图-策略金字塔模型

意图分析与转译并非单向过程，而是与网络状态形成动态反馈。当转译过程中发现资源不足或约束无法满足时，系统会反向反馈至意图获取环节，请求用户调整意图或补充约束条件。这种闭环反馈确保了意图与网络能力的匹配，避免无效操作。

（三）策略验证

策略验证的核心目标是在策略下发至实际网络前，通过系统性校验确保策略的可行性、正确性及安全性，并满足用户意图的预期效果。从验证目标来看，策略验证需覆盖四个核心维度：首先是策略与意图的一致性，即校验转译生成的策略是否准确映射用户原始意图，避免因转译偏差导致目标偏离。其次是资源与约束的可行性，即检查策略所需的网络资源是否充足，以及是否满足时间、地理等约束条件。接着是策略的冲突检测，识别并解决策略间的矛盾。最后是安全性验证，即验证策略是否存在潜在风险，确保策略执行后不会影响网络整体稳定性。

在技术实现上，策略验证主要依赖数字孪生与形式化验证两种核心方法。数字孪生技术通过构建与物理网络实时同步的虚拟镜像，将待验证策略注入虚拟环

境中模拟执行，全方位监测网络状态变化。形式化验证则通过数学建模将策略转化为逻辑表达式，利用定理证明、模型检测等工具校验策略的正确性。在实际应用中，两种方法常结合使用。验证完成后需将结果反馈至上游模块：若策略通过验证，则允许策略进入下发流程；若存在问题，则明确标注错误类型及建议调整方向。

策略验证通过系统化的技术手段，在策略执行前全面排查潜在风险，确保意图转化的准确性与网络运行的稳定性。随着网络异构性与业务复杂性的提升，策略验证将进一步融合人工智能技术，实现从被动验证到主动预防的升级。

（四）意图下发与执行

意图下发与执行是指将经过验证的网络策略精准高效地部署到网络中，并通过实时监控确保意图的最终落地。在意图下发阶段，系统首先需完成意图的分层传递与域间协同。由于网络通常由多个域构成，而单一意图可能涉及跨域资源调度，因此下发过程需遵循“分层解耦、协同联动”的原则。下发过程中，系统需具备柔性调度能力，以适应网络的动态变化。同时，下发过程支持优先级机制，对于高紧急度意图，可中断低优先级意图的执行资源，实现秒级响应。

意图执行阶段的核心是将下发的策略转化为网络设备的具体配置，并通过实时感知确保执行效果。为确保执行的准确性，系统引入双闭环监控机制：内层闭环聚焦单域执行效果，由域内控制器实时校验设备配置与策略的一致性；外层闭环则关注端到端意图达成情况，通过跨域数据聚合，判断整体意图是否满足。若发现偏差，系统会自动触发局部优化，无需人工干预。在复杂场景中，意图执行还需具备自适应演进能力。系统通过内生感知能力动态调整执行策略，无需用户重新输入意图。

（五）意图的实时反馈

意图的实时反馈是指通过持续采集、分析网络运行数据，将意图执行后的实际效果与预期目标进行比对，并将结果实时反馈至上游模块，为意图的动态调整提供依据。其核心目标包括：

- 验证意图执行效果：通过实时数据确认策略是否实现了用户意图。
- 动态调整策略：当网络状态与意图要求不符时，触发策略修正或资源重新分配。
- 提升网络自愈能力：通过闭环反馈机制，使网络具备自动检测故障、预测风险并快速恢复的能力。
- 优化资源利用率：基于实时数据动态调整资源分配，避免资源浪费或不足。

实时反馈涵盖三类关键信息：一是业务指标达成情况，即直接反映意图目标的核心参数。这些指标通过部署在网络节点的传感器、探针或内置监控工具实时采集。二是网络资源状态，包括链路负载、设备 CPU 利用率、内存占用、能源消耗等，用于评估策略执行对网络整体资源的影响。三是异常事件与故障告警，如链路中断、设备故障、安全攻击等突发情况，这些信息需以最高优先级上报，确保系统快速响应以避免意图失效。

7.2 数字孪生网络

7.2.1 数字孪生网络的定义和架构

（一）数字孪生网络的定义

数字孪生网络是以数字化方式创建物理网络实体的虚拟孪生体，通过实时数据交互实现虚实映射，支持网络全生命周期管理的智能化网络系统。在该系统中，各种网络管理和应用可利用数字孪生技术构建网络虚拟孪生体，基于数据和模型

对物理网络进行高效的分析、诊断、仿真和控制，助力网络实现低成本试错、智能化决策和高效率创新，实现网络的高水平自治。数字孪生网络包括数据、模型、映射和交互四大核心要素。

图 24 数字孪生网络核心要素

- 数据为整个系统提供底层支撑，采集物理网络的设备参数、拓扑关系、运行状态等全量数据，为模型构建、状态评估和决策分析提供依据。

- 模型负责构建虚拟网络的数字化表示，为模拟网络元素和资源的配置、状态或使用变化提供了基础。

- 映射是数字孪生网络实现虚实对应的关键机制，确保虚拟孪生体与物理网络在结构、状态和行为上的高度一致性，可替代物理网络进行仿真、推演与决策。

- 交互是连接物理网络与虚拟孪生体的桥梁，实现物理网络与虚拟孪生体之间的实时数据传输与控制指令下发，确保虚实状态同步。

数字孪生技术中常提到物理实体和数字孪生体的概念，数字规划体则是基于网络自治特点提出的新概念。物理实体是数字孪生网络映射的对象，指构成通信

网络的所有物理基础设施及相关要素，是虚拟空间建模的原型。对于硬件而言，就是硬件形态本身；对于软件而言，是软件的载体，如镜像文件、软件代码等。

数字孪生体是物理实体在虚拟空间的数字化复现，通过数据建模与实时映射，精准反映物理实体的配置、状态及行为规律，与物理实体状态保持同步，是数字孪生网络的核心载体。

数字规划体是基于物理实体当前状态与未来目标，通过智能化规划生成的前瞻性数字化模型，是网络向自治演进的核心规划工具。它并非是对物理实体现有状态的简单映射，而是聚焦未来时刻的优化目标，融合业务需求、资源约束与网络能力，构建物理实体应达成的理想状态参数集，为网络迭代提供可量化、可执行的路径指引。

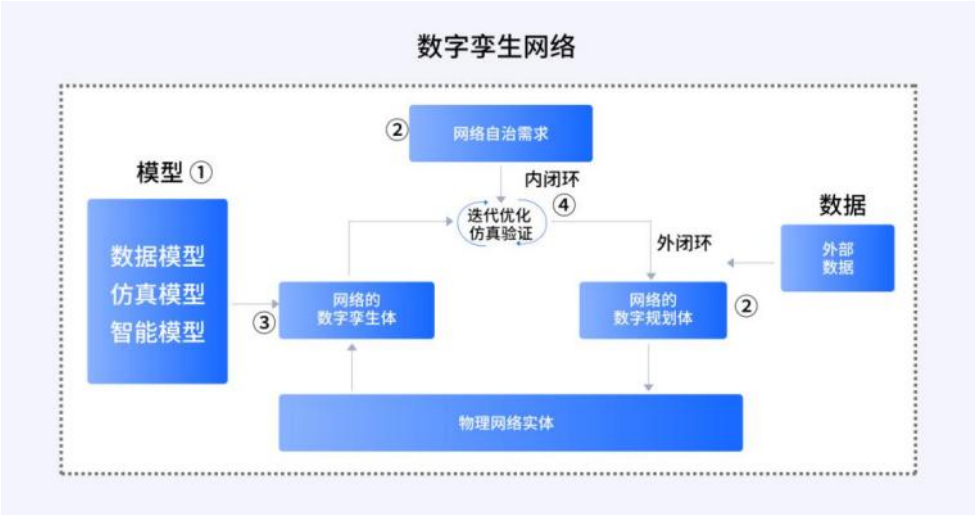

图 25 数字孪生网络基本概念间的关系

（二）数字孪生网络架构

结合数字孪生的技术特点和通信网络的需求，中国移动提出了“三层三域双闭环”数字孪生网络参考架构。“三层”指构成数字孪生网络系统的物理网络层、孪生网络层和网络应用层；“三域”指孪生网络层数据域、模型域和管理域；“双

闭环”是指孪生网络层内基于服务映射模型的“内闭环”仿真和优化，以及基于三层架构的“外闭环”对网络应用的控制、反馈和优化。

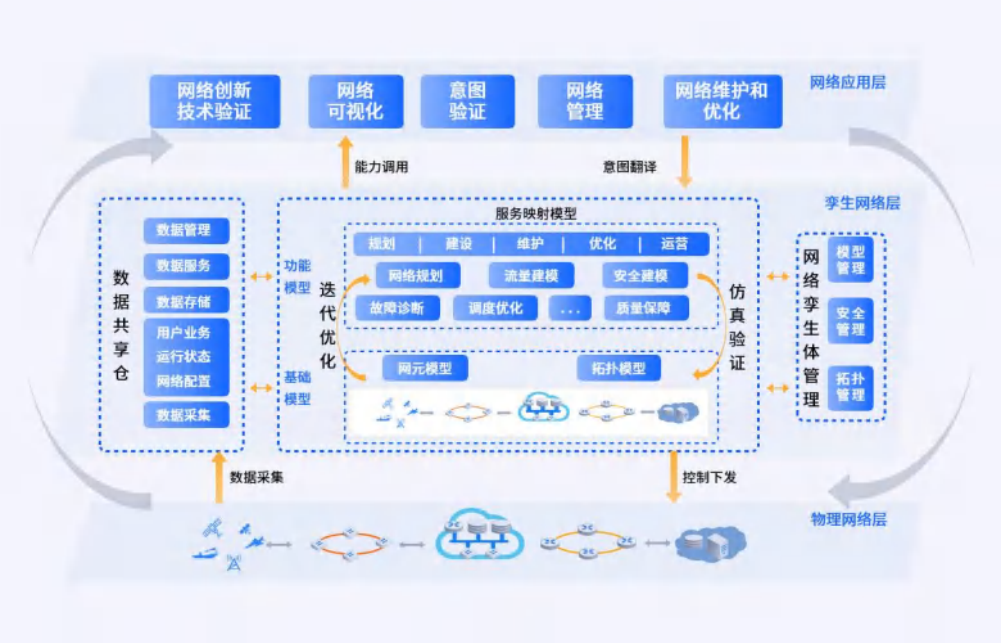

图 26 数字孪生网络架构

(1) 三层

物理网络层是数字孪生的映射对象，由构成端到端网络的所有物理实体组成，包括基站、交换机、路由器、终端设备、配套设施及环境要素。其核心功能是通过南向接口向孪生网络层传输实时数据，同时接收孪生网络层下发的控制指令，实现与虚拟空间的基础交互。

孪生网络层是数字孪生网络的核心，负责构建物理网络的虚拟镜像并实现智能化管理，包含数据共享仓库、服务映射模型和网络孪生体管理三个关键子系统。

网络应用层面向用户或业务需求，通过北向接口向孪生网络层输入需求，并通过模型化实例在孪生网络层进行业务的配置。充分验证后，孪生网络层通过南向接口将控制更新下发至物理实体网络。

(2) 三域

三域是孪生网络层的内部构成，实现数据、模型与管理的协同。数据域对应数据共享仓库子系统，负责采集和存储物理网络的多源异构数据，为各种服务于应用的网络模型提供准确完备的数据。模型域对应服务映射模型子系统，基于数据构建多样化模型，包括基础模型和功能模型，向上层网络应用提供服务，最大化网络业务的敏捷性和可编程性。管理域对应孪生体管理子系统，负责虚拟孪生体的全生命周期管控，确保虚拟孪生体与物理网络的长期一致性。

（3）双闭环

双闭环是数字孪生网络实现智能化的核心机制，确保虚拟仿真与物理执行的协同优化。

内闭环在孪生网络层内部，通过服务映射模型对网络策略（如参数调整、路由优化）进行仿真验证，评估可行性并迭代优化，避免直接作用于物理网络导致的风险。例如，在网络切片优化中，先在虚拟空间仿真不同切片的带宽分配方案，验证其对 SLA 的达成能力。

外闭环将内闭环验证后的最优策略通过南向接口下发至物理网络，监测物理网络的执行效果并反馈至孪生网络层，进一步修正模型与策略，形成“仿真-执行-反馈-优化”的循环。例如，在网络智能容灾中，孪生网络层生成的倒换策略经物理网络执行后，将倒换耗时、业务中断率等数据反馈至虚拟空间，持续优化容灾模型。

7.2.2 数字孪生网络的关键技术

（一）全景数据服务技术

全景数据服务技术是数字孪生网络的数据基础，旨在实现物理网络全量数据的精准采集、整合与服务，为模型构建与决策分析提供支撑。其核心包括三个层面：

- 多源异构数据采集：通过网络遥测、传感器接入、协议解析等技术，采集物理网络的设备参数、运行状态、环境数据及业务数据。

- 异构数据整合与存储：针对结构化、半结构化、非结构化数据，采用分布式存储技术构建统一数据仓库，通过抽取、转换、加载流程实现数据清洗、去重与标准化。

- 统一数据服务接口：提供标准化的数据查询、调用与共享接口，支撑上层模型对数据的按需获取，同时保障数据访问的安全性与高效性。

（二）网络全生命周期建模技术

建模技术是数字孪生网络的核心引擎，通过抽象物理网络的静态特征与动态行为，构建可复用可组合的模型体系，分为以网络设备基本配置、环境信息、运行状态、链路拓扑等为代表的基础模型和以网络感知、分析、仿真、推理、决策等为代表的功能模型两大类，通过模型编排技术将基础模型与功能模型灵活组合，满足复杂场景需求，同时基于实时数据持续迭代模型参数，确保与物理网络的长期一致性。

（三）全域孪生体管理技术

全域孪生体管理技术负责虚拟孪生体的全生命周期管控，确保其与物理网络的动态适配与精准映射，核心包括：

- 数字线程整合：通过贯穿网络规划、建设、运维、优化全流程的数字线程，整合各阶段数据，实现孪生体的溯源。

- 动态更新与偏差修正：基于实时采集的物理网络数据，持续校准孪生体模型参数，当物理网络发生变化时，自动触发孪生体重构。

- 多粒度孪生体管控：支持从单网元到端到端网络的多粒度孪生体创建与管理，通过层级化管控实现资源高效调度。

（四）网络可视化技术

通过可视化与交互技术，实现虚拟孪生体对物理网络的直观呈现与动态操控，核心包括：

- 多维度可视化：基于 2D/3D 建模、GIS、BIM 等技术，呈现网络拓扑、设备状态、流量分布与故障位置。
- 沉浸式交互与仿真：支持用户通过图形界面、VR/AR 等方式与孪生体交互，如拖拽式调整网络参数、模拟故障场景。
- 关联分析与态势感知：通过多维度数据关联，挖掘隐藏规律，生成网络态势评估报告。

（五）全向接口协议技术

全向接口协议技术通过标准化接口实现物理网络、孪生网络与应用层的高效交互，确保数据传输与控制指令的可靠性与实时性，包括三类接口：

- 南向接口：连接物理网络与孪生网络层，支持数据采集与控制指令下发，采用 RDMA 等技术降低传输时延。
- 北向接口：连接孪生网络层与应用层，提供模型调用、数据查询等能力，采用基于 QUIC 的 HTTP/3.0 协议。
- 内部接口：支撑孪生网络层内数据域、模型域、管理域的协同，采用高效序列化协议确保数据交互效率。

7.2.3 基于 DTN 实现意图驱动的网络

数字孪生网络为意图驱动网络的有效部署方面提供了坚实基础，它能够实现网络配置的预先验证以及用户业务意图的实时保障等关键功能。物理网络层作为实体基础，通过实时数据采集将物理世界状态同步至上层。孪生网络层作为智能

中枢，整合共享数据仓库中的多源信息，并依托服务映射模型实现核心功能。网络应用层直接对接用户，接收业务意图输入。

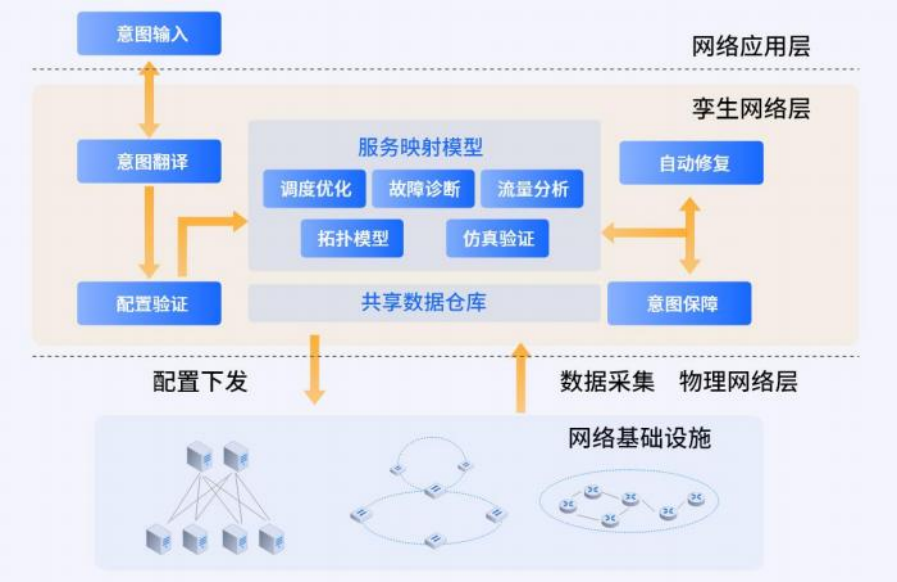

图 27 基于 DTN 实现 IDN 的参考框架

在具体运作中，数字孪生网络通过两大机制保障意图的可靠落地：

（一）意图驱动的预验证机制

用户意图经转译生成网络配置命令后，并非直接下发至物理网络。因为物理网络承载着多业务运行，直接下发配置可能引发地址冲突、路由环路等风险，影响既有业务。此时，孪生网络层的服务映射模型发挥预演作用，在虚拟环境中模拟配置下发后的网络行为，同时检测对其他业务的潜在干扰。验证通过后，配置才会安全下发至物理网络。这样不仅可以确保新配置满足用户业务需求，又不会对现有网络服务产生负面影响，从源头规避试错成本。

（二）意图偏离的自修复机制

孪生网络层持续从共享数据仓库获取物理网络的实时状态，服务映射模型不断校验用户意图的达成情况。当监测到网络偏离意图，借助故障诊断、AI 根因分析等能力，定位问题源头并生成修复策略。但修复策略直接下发物理网络风险较高，此时需再次依托服务映射模型，对修复策略进行仿真验证：利用仿真验证

模拟策略执行后的网络状态，确认能否有效解决问题且不引发新故障。通过验证的策略，由自动化配置模块下发至物理网络，实现故障自动修复，既摆脱人工确认的低效依赖，又推动 AI 技术安全落地，提升运维效率与网络自治水平。

7.3 智能网络大模型

7.3.1 智能网络大模型的核心应用

大模型凭借强大的学习能力和泛化性能，成为推动网络智能化升级的核心动力之一。它能深度挖掘网络数据中的潜在价值，精准把握网络复杂多变的运行状态，在高效故障诊断与预测、网络资源优化配置等关键领域发挥重要作用，为网络运维、管理及性能优化等提供智能决策支持，进而改善网络服务质量，提高用户体验。

（一）网络智能运维

大模型能够学习复杂的模式，并且自动识别异常行为。在运维场景中，日志分析、系统指标分析、本机调用链分析等涉及的数据多为非结构化或时序数据，这类数据的特征挖掘与异常识别非常适合采用深度学习模型。通过将实时采集的网络数据输入到训练完备的大模型中，模型能依据学习到的模式和规律，对当前网络状态进行精准评估与判断。一旦检测到异常数据模式，会立即发出故障预警信号，并初步判定故障类型及可能位置。大模型用于异常检测的基本思路可以概括为以下几种方式：

- 日志异常检测：借助 BERT 等 NLP 预训练模型学习正常日志的语义与格式模式，进而识别偏离正常模式的异常日志。

- 时间序列预测：利用 Transformer 等模型对系统指标的变化趋势进行预测，当实际指标与预测结果的偏差超过阈值时，判定为异常。

- 无监督学习：通过自编码器（Autoencoder）、对比学习（Contrastive Learning）等方法，在无标注数据的情况下挖掘数据分布规律，从而检测未知异常。

与传统单一指标检测方法相比，大模型能够综合多维度数据信息开展故障检测，大幅提升了检测的全面性与准确性。不仅如此，在检测到故障后，大模型还可凭借对网络数据的深度理解与推理能力，深入分析故障相关数据，通过追溯网络事件的时序关联、剖析相关设备的配置信息及运行状态变化，精准定位故障的根本原因，为快速解决故障奠定基础。

（二）网络性能优化

大模型凭借其强大的多模态数据分析处理能力和动态推理能力，正在从资源调度、流量管控、拥塞控制等多个维度重塑网络性能优化的边界。

传统资源分配依赖固定规则，而大模型支持动态优化配置，能依据实时网络状态自适应调整。凭借动态感知能力，大模型可实现跨域资源全局统筹，在保障核心业务服务质量的同时，灵活调整非关键业务资源占比，避免局部拥塞与闲置。同时，它能随网络负载实时变化持续优化分配策略，将冗余资源导向需求节点，确保网络整体效率最优，使资源供给与业务需求始终动态平衡，从根本上提升资源利用效率与响应速度。

在流量管控中，大模型通过对网络流量的实时深度解析与动态趋势预判，构建起智能化的流量调度体系。它能够全面捕捉流量的来源、类型、传输路径及负载特征，结合网络拓扑结构与链路状态，精准识别流量的正常波动与异常增长，形成对全网流量的全局感知。基于这种感知能力，大模型可以根据不同流量的业务优先级与服务质量需求，制定差异化的管控策略。这种动态适配的管控模式，不仅能提升网络带宽的整体利用率，还能确保各类流量在复杂网络环境中始终保持高效、稳定的传输状态，为网络服务质量提供坚实保障。

（三）网络安全防护

在网络安全防护中，大模型依托其多模态数据处理、模式学习及动态决策等特性，可实现对复杂威胁的深度感知与快速处置，构建起从威胁识别到主动防御的智能化体系。大模型可同时处理网络流量、终端日志、威胁情报、用户行为等多源异构数据，通过特征工程优化与深度学习建模，提取正常行为与异常行为的差异化特征。

大模型通过持续学习新的攻击样本，能够捕捉威胁的本质特征。对于已知威胁，模型会比对实时数据与知识库中的特征，实现毫秒级识别；对于未知威胁，模型通过无监督学习分析正常行为基线，当出现偏离基线的异常时，会标记为潜在威胁，并通过关联分析判断威胁等级。识别威胁后，大模型会根据威胁类型、影响范围和目标重要性，自动生成差异化防御策略，并联动多个安全设备执行。

最后，大模型会实时监测防御措施的执行效果，若防御未达预期，模型会回溯分析漏洞，随后重新学习该变种样本的特征，更新识别模型，并优化防御策略。同时，模型会将每次攻防案例纳入训练数据，不断丰富威胁知识库，使防御能力随攻击手段的进化而同步提升。

7.3.2 多智能体（Multi-Agent）群智协同

智能体（Agent）是一种能够实时感知环境、自主决策并采取行动以实现特定目标的智能系统，它是将大模型能力充分发挥出来的重要手段，通过不同智能体之间的相互协同可以助力解决通信网络的复杂问题，在高阶自智场景中发挥更大价值。AI Agent 由规划（Planning）、记忆（Memory）、工具（Tools）与行动（Action）四大关键部分组成，分别负责任务拆解与策略评估、信息存储与回忆、环境感知与决策辅助、以及将思维转化为实际行动。

图 28 AI Agent 应用的基础技术架构

规划模块负责将网络目标拆解为可执行的子任务，并评估策略的可行性与潜在风险；记忆模块通过存储网络历史数据、拓扑关系、故障模式等信息，为决策提供经验支撑；工具模块整合网络感知、数据分析、协议解析等能力，将网络状态转化为智能体可理解的信号；行动模块则通过标准化接口将决策转化为具体操作，并通过反馈机制动态调整。

从工作模式来看，智能体可以分为单智能体、多智能体和混合智能体。单智能体只有一个智能体进行感知、学习和行动，与环境独立交互，根据环境反馈优化下一步行动策略以实现预期目标。多智能体是一种特殊的智能体，每个智能体都有自己的感知、决策和行动能力，并与其他智能体进行交互协作和信息共享，共同实现复杂的目标。混合智能体是由智能体和人类共同参与决策过程，强调人机协作的重要性和互补性。

多智能体群智协同作为一种先进的技术范式，由多个独立自主的智能体在网络中进行无缝协作，其核心在于通过分布式协作机制实现对复杂网络的动态适配与高效管理，这种协作模式与网络的分布式架构、异构特性深度契合。在功能架构上，多智能体采用分层协同设计，不同智能体被赋予特定领域的专精能力，例如负责实时流量监测的感知智能体、专注资源调度的决策智能体、承担安全防护的防御智能体等，各智能体通过明确的角色定位形成功能互补。

在此基础上，协作模式强调动态任务拆解与目标对齐，当面临全局网络优化、跨域故障处理等复杂任务时，由负责统筹的协调智能体将任务分解为子目标，依据各智能体的能力边界与实时负载进行分配，同时通过持续的状态同步确保子目标与全局目标的一致性，避免局部优化对整体网络性能的干扰。此外，协作模式还包含自适应的交互机制，智能体之间通过标准化的通信协议交换实时数据与决策意图，既可以是集中式的信息汇聚，也支持分布式点对点交互。

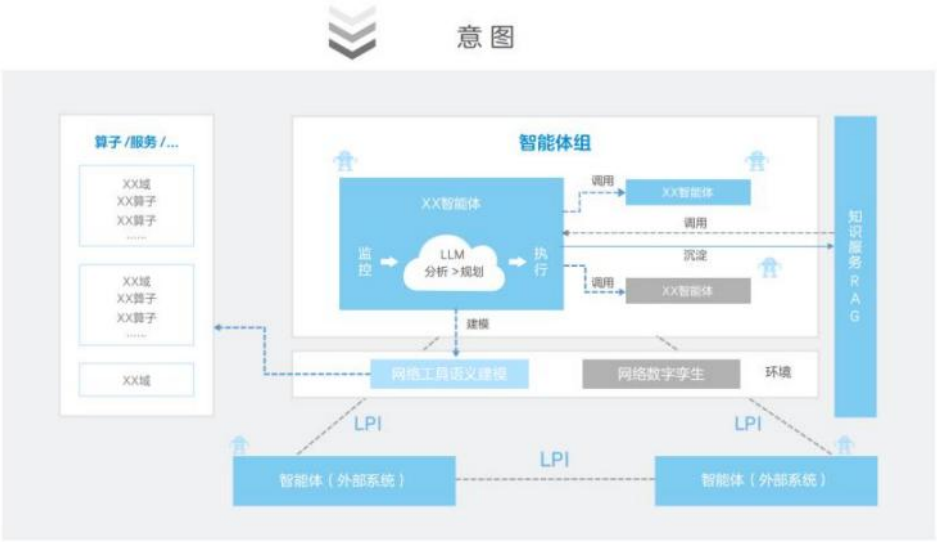

图 29 多智能体群智协同

7.3.3 Agentic SOAR 智能化网络安全编排架构

智能化安全编排和自动化响应（Agentic SOAR，Agentic Security Orchestration、Automation and Response）是以 AI Agent 为核心驱动，将 LLM 的动态推理能力与 SOAR 的自动化执行能力深度融合的新型安全运营架构。其核心目标是打破传统 SOAR 静态依赖的局限，实现从流程自动化到目标自主化的范式升级，解决安全运营中告警疲劳、响应滞后、专家依赖等痛点。

Agentic SOAR 的运作模式与传统 SOAR 存在明显差异。传统 SOAR 遵循的是一条从手动设计到静态响应的线性路径，而 Agentic SOAR 则是以目标意图为起点，由 LLM Agent 进行动态循环的推理、规划、执行过程。

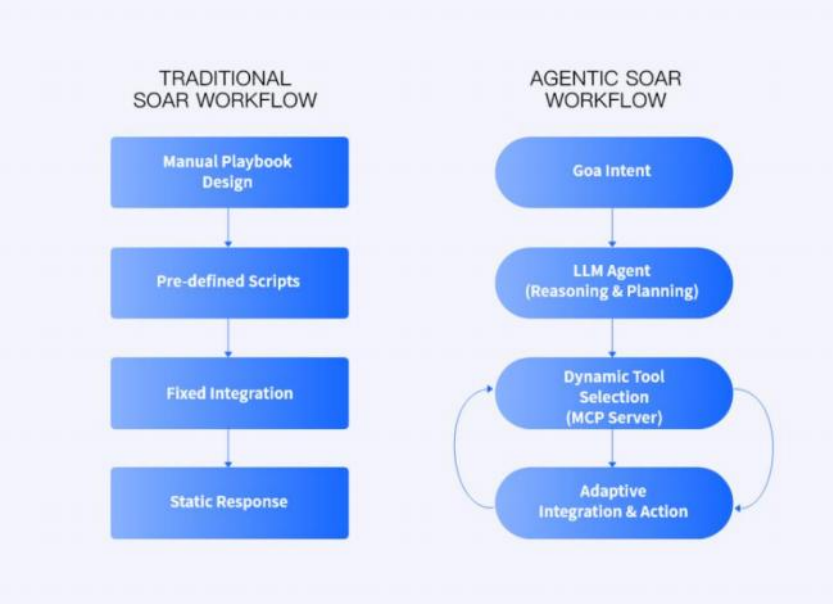

图 30 传统 SOAR 与 Agentic SOAR 工作流对比

目标意图设定（Goal Intent）：Agentic SOAR 工作流程始于明确目标意图，与传统 SOAR 依赖人工手动设计模式不同，Agentic SOAR 中的目标意图无需预先设定详细的操作步骤，它更侧重于明确最终想要达成的结果，为后续流程提供方向指引。

LLM Agent 推理与规划（Reasoning & Planning）：明确目标意图后，系统会调用 LLM Agent 进行推理与规划。LLM Agent 基于大语言模型强大的自然语言理解和生成能力，对目标意图进行深入解析。它会分析当前任务的特点、相关的历史数据以及可能存在的约束条件等信息，进而制定出一套详细的应对策略，包括需要采取的具体操作、操作的先后顺序以及可能的风险应对措施等。

动态工具选择（Dynamic Tool Selection - MCP Server）：完成推理与规划后，LLM Agent 的决策结果会被传递到动态工具选择环节。这里的 MCP Server（多组件平台服务器）类似于一个工具资源池，存储着各类不同功能的工具。LLM Agent 根据制定的应对策略，从 MCP Server 中动态选择最适合当前任务需求的工具。与传统 SOAR 中固定的工具集成方式不同，Agentic SOAR 的动态工具选择能够根据实时任务需求灵活调配资源，提高了系统应对复杂多变场景的能力。

自适应集成与行动（Adaptive Integration & Action）：在这个阶段，所选择的工具会被自适应地集成到工作流程中，并按照 LLM Agent 规划的策略执行相应操作。在操作执行过程中，系统会实时监控执行的效果和状态，并将反馈信息及时传递给 LLM Agent。如果执行结果未达到预期，LLM Agent 可以根据反馈信息重新进行推理和规划，调整工具选择和操作策略，形成一个闭环的自适应调整机制，确保最终能够实现最初设定的目标意图。

7.4 联邦学习

7.4.1 联邦学习的定义

在网络智能化升级中，数据是核心驱动力。然而，数据的采集、处理和应用过程中往往涉及用户隐私问题，如何在利用多方数据进行智能分析的同时，有效保护数据隐私和安全是一个亟待解决的问题。联邦学习（FL，Federated Learning）技术能够在不泄露数据隐私的前提下，打破数据孤岛，解决数据传输瓶颈，赋予网络新智慧。

在网络场景中，数据天然分布于边缘设备、基站、核心网节点等不同位置，且包含用户行为、设备状态、业务交互等敏感信息，直接集中数据进行模型训练既面临传输成本过高的问题，也存在隐私泄露风险。联邦学习是一种分布式机器学习框架，允许各节点在不共享原始数据的前提下在本地基于自有数据完成模型训练，通过加密传输模型参数或梯度信息进行全局协同优化。其核心思想是“数据不动模型动”，即数据保留在本地，仅通过交互中间结果实现协作学习。既避免了原始数据的暴露，又能让全局模型聚合各节点的局部知识，为网络智能化提供了数据安全基础。

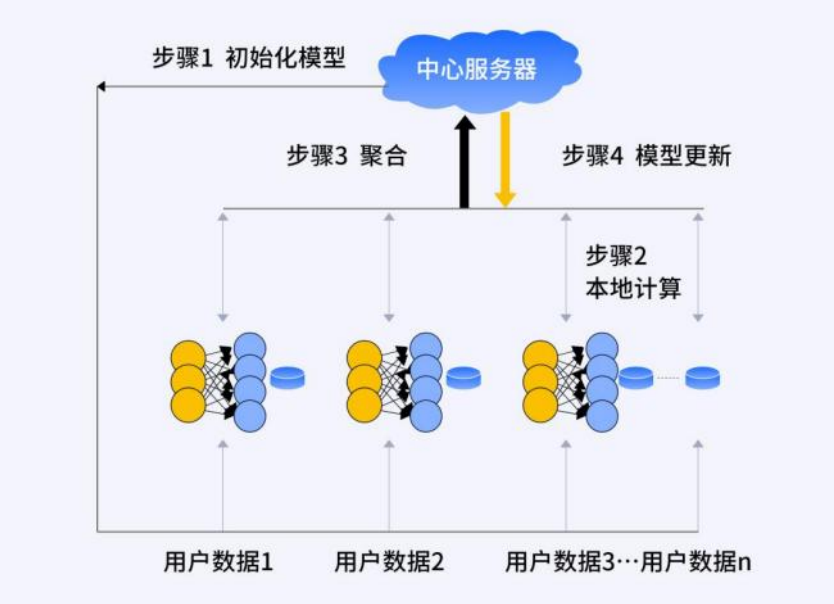

图 31 用户－服务器学习流程

联邦学习在跨域分析中的实现依赖于一个闭环的迭代流程，即“初始化-本地训练-全局聚合-模型下发”。

- 步骤 1 初始化阶段：中央服务器初始化一个全局模型，并将模型参数分发给各个参与方。

- 步骤 2 本地训练阶段：每个参与方使用本地数据集对收到的全局模型进行训练，生成本地模型更新。在这个过程中，原始数据始终保留在本地，不进行交换，从而保护了数据隐私。

- 步骤 3 全局模型聚合阶段：参与方将本地模型更新（如梯度、模型参数等）加密后上传至中央服务器，加密技术确保了数据在传输过程中的安全性。中央服务器收集所有参与方的模型更新，并进行聚合，生成新的全局模型。聚合过程旨在结合各方模型的优势，提高全局模型的性能。

- 步骤 4 模型下发阶段：更新后的模型下发至各节点进行下一轮训练，重复步骤 2 至步骤 4，直到全局模型达到满意的性能或收敛。通过多轮迭代训练，全局模型不断优化，以适应更广泛的数据分布。

7.4.2 联邦学习的分类与关键技术

根据联邦学习的数据特点，即不同参与方之间的数据重叠程度，联邦学习主要分为横向联邦学习、纵向联邦学习和联邦迁移学习。

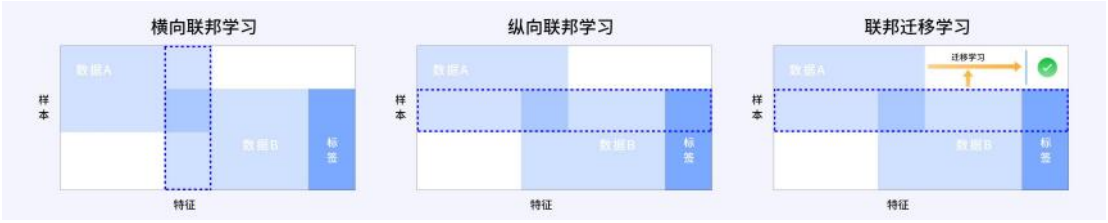

图 32 联邦学习分类

横向联邦学习适用于参与方数据特征重叠较多但样本重叠较少的情况。例如，不同地区的银行拥有相似的业务但不同的客户群体，它们可以通过横向联邦学习联合训练模型，提高模型的泛化能力。

纵向联邦学习适用于参与方数据样本重叠较多但特征重叠较少的情况。例如，同一地区的银行和电商拥有相似的客户群体但不同的数据特征，它们可以通过纵向联邦学习联合训练模型，实现特征互补。

联邦迁移学习适用于参与方数据样本和特征重叠都很少的情况，它将联邦学习的概念加以推广，可在任何数据分布、任何实体上进行协同建模以学习全局模型。

在联邦学习中，安全性与效率的平衡依赖于差分隐私、同态加密、安全多方计算等核心技术的协同支撑。

差分隐私通过在局部模型参数或梯度中引入精心设计的噪声，模糊个体数据对全局模型的影响，从而阻止攻击者通过逆向工程从聚合参数中反推原始数据信息。其核心在于控制噪声的强度，既需确保噪声足够掩盖单一个体的数据特征，又要避免过度噪声导致模型精度下降，这种平衡通过隐私预算机制实现，每个参

与节点的参数更新被分配一定的隐私预算，累计消耗不超过阈值，既保障了数据隐私，又为模型收敛提供了基础。

同态加密技术则聚焦于参数传输与聚合过程的加密保护，其独特之处在于支持对加密数据直接进行加法、乘法等数学运算，无需解密即可完成模型参数的聚合计算。在联邦学习的整个过程中，原始参数始终处于加密状态，服务器与其他节点均无法窥探个体数据，从根本上杜绝了传输与聚合环节的隐私泄露风险。

安全多方计算通过密码学协议将模型参数的聚合过程转化为分布式协同计算，使多个参与方在不泄露各自私有数据的前提下，共同完成全局参数的计算。其核心逻辑是秘密共享，每个节点将本地参数拆分为多个份额，分别发送给其他参与方，任何单一参与方仅持有部分份额，无法还原完整参数；聚合时，各参与方基于所持份额进行局部计算，再通过协议组合结果，最终得到全局聚合参数，而整个过程中没有任何一方能获取其他节点的原始参数或完整份额。

第 8 章　AI for Network 典型应用实践

本章聚焦 AI 技术在网络领域的落地成果，精选部分 AI for Network 应用案例，深入剖析其技术架构、实施路径与价值创造，为行业提供可借鉴的实践经验。

8.1　中国联通 AI 智能体助力地铁无线网优创新

为解决地铁场景人工测试耗时耗力的诸多问题，中国联通提出了一种新的地铁用户识别和定位方法，依靠地铁用户上报 MR 来实现替代人工测试模式，并嵌入网络专家大模型，自动进行地铁问题识别、根因分析、智能决策，并对执行效果进行自动闭环评估，从而实现了地铁网络问题的全程管控。“大模型中枢+小模型执行”架构，推动地铁无线网络优化从人工经验驱动升级为 AI 目标驱动，实现降本、提质、增效三重突破。

该方案通过大数据分析用户连续通过同方向站台或轨道区间的特征，提升地铁用户及起止时间识别准确性；收集地铁指纹数据，结合切换锚点思路将 MR 和信令数据地理化，助力精准定位网络问题。基于全国现网数据，构建无线网优专家多智能体，支持自然语言与图形化双模式操作，减少人工依赖，自动生成解决方案。借助 Chat 交互打破操作限制，实现远程自动路测和问题统计，提升效率与用户体验，同时节省成本。

图 33　网络 AI 大模型应用设计图

该方案有三大创新：一是网络 AI 大模型应用创新，打造质差分析专家知识库，结合大模型意图识别、智能体任务规划与 RAG 技术分解任务，利用 DeepSeek 推理提供优化建议；二是构建地铁用户行为精准识别模型，采集 XDR 信令和 MR 数据建模，聚类分析特征以识别用户，监控评估网络质量与用户感知；三是基于地铁用户精准定位模型，研究隧道和轻轨定位算法，构建定位指纹库，建立智能定位模型，实现地理化分析。

8.2 中国移动九天大模型助力无线网络优化智能升级

中国移动九天网络大模型充分发挥无线网络运行数据时空关系的复杂性，精准解决无线网络日常运维作业中对专家水平要求高、工单处理时效差、问题分析维度不足、优化效率低等痛点。九天网络大模型以其卓越的内生能力，可嵌入到无线网日常优化中，提供问题诊断、分析决策、方案执行、效果预估的能力，实现对传统无线网优化流程的智能化重塑，打造无线问题端到端的自闭环优化。

图 34 九天·网络大模型助力无线网络优化升级

依托九天·网络大模型卓越的内生能力优势，结合无线网络结构化数据为主的特点，AI+无线网络自优化应用可嵌入到无线网运维一线生产作业中，提供四大核心功能：

- 智能感知：相较于基于无线接通率等传统核心指标制定的劣化规则，大模型能结合更多维度的网络日志准确分辨网络波动和网络异常，计划实现更贴合用户真实感知的无线工单挖掘；

- 智能诊断：大模型面向全场景无线性能工单快速确定问题根因，无需人工对疑点逐一排查；

- 智能决策：大模型根据派单信息拉取多网元性能数据，一步生成优化方案，可支持参数类、硬件资源类的优化方案输出；

- 智能预测：大模型结合小区历史数据学习情况对方案执行后的效果进行下发前评估，针对达不到预期效果的方案终止实施，减少对网络的影响。

8.3 中国铁塔网络智能化运维与优化平台

为解决用户在网络运维与优化方面的需求与痛点，中国铁塔南京科技创新中心深入总结当前运维工作需求，构建了基于 SDN/NFV、AI 大模型的网络智能化运维与优化平台。平台采用分层架构设计，涵盖数据采集、处理、分析、决策及交互五大层级。

图 35 平台实现运维智能化升级

平台具备三大核心优势：

- 全域设备统一纳管：通过多协议纳管引擎实现华为/H3C 等 200+型号设备的统一接入。通过 CMDB 构建多层拓扑与 GIS 视图，形成全息设备画像，实现从基础设施到业务系统的全生命周期数字化管理。

- 多维度实时监控：通过基础设施与业务级双维度实时监控体系，实现全网状态精准感知。在基础设施层，采用容器化探针对服务器、基站及网络设备的全量指标进行秒级采样；在业务层，通过 SRv6 路径探针实现核心网业务流的毫秒级追踪，动态监测时延、丢包率等关键指标。

- 智能告警与自愈：告警闭环管理系统通过 LSTM 模型实现告警多维收敛，并采用三级分级推送机制（严重/重要/提示）联动短信/邮件/企微等渠道；同时自动关联知识库生成修复工单，结合 100+预置运维剧本实现 ms 级自动化响应，并通过虚拟机热迁移等技术确保设备自愈。

核心功能涵盖四方面：一是监控与高效利用，包括网络资源灵活调度、基站需求预测及边缘节点资源共享；二是网络性能监测与优化，涉及实时指标监控、智能路由调整及流量管理；三是 AI 驱动的智能监测与分析，实现异常检测、预测性维护及智能决策支持；四是一体化监测与管理，提供监控总览、全域资源监控、网络智能编排及告警管理等功能。

8.4 华为星河 AI 网络解决方案

华为星河 AI 网络通过全系设备的 AI 加持与赋能，与 NetMaster 网络智能体深度协同，实现了更精准的业务感知、更实时的业务闭环，从根本上重塑网络的体验、安全和运维。华为星河 AI 网络覆盖四大领域：

- 星河 AI 数据中心网络：通算场景打造高韧性数据中心网络，智算场景构建高算效数据中心网络。

- 星河 AI 广域网：通过全业务融合、确定性体验保障和最佳运维体验，聚焦融合、体验、智能，打造高运力融合广域网，构筑智能融合 IP 网络底座。

- 星河 AI 园区网络：通过无线体验升级、应用体验升级、运维体验升级和安全体验升级，牵引 AI 使能以体验为中心建网，跃升企业数智生产力。

- 星河 AI 网络安全：通过 AI 赋能网络安全防护能力，打造智检测、智联动、智融合的 SASE 解决方案，构筑企业网络安全防御防线。

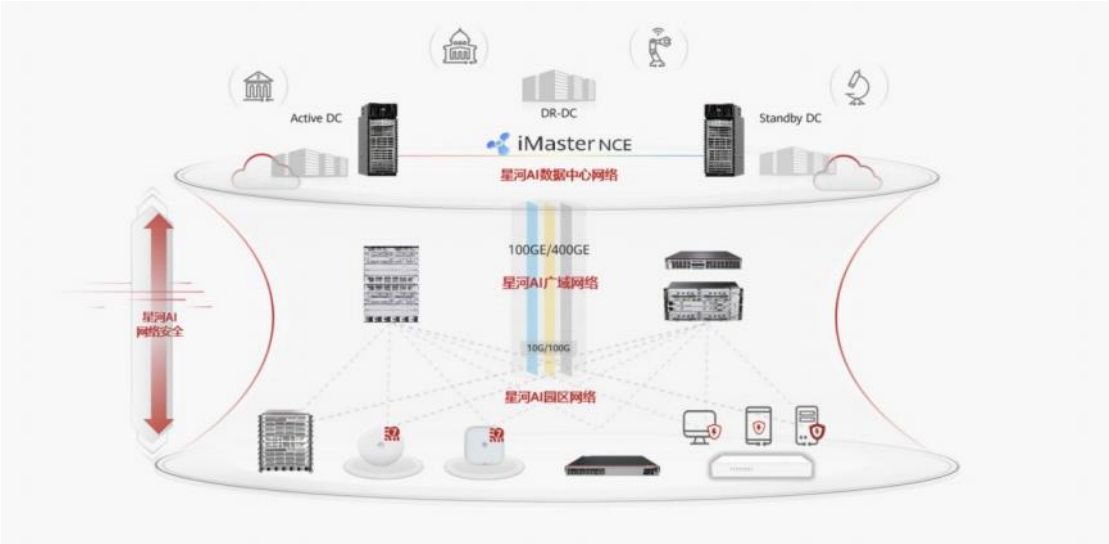

图 36 华为星河 AI 网络解决方案

其中，华为星河 AI 网络智能体 NetMaster 集成了华为数据通信领域上千亿语料，以及 1 万多名网络专家的经验，具备强大的语义理解能力，是通信网络领域的重要突破。其支持运维数据问答、交互式业务分析与辅助决策，通过独家 AI 网络思维链，实现典型场景下 80%无线故障自诊断、自动生成处置建议并自动执行，从而推动网络迈向智能化新时代。

iMaster NCE 网络数字地图以类似交通导航地图的方式，将网络空间和物理空间进行深度融合，基于数字孪生理念构建企业数字化智能大脑，实现云、网、端、应用及用户的统一智能管理。实现从网络到应用的全息可视，并提供了智能路径导航与智能优化等能力，显著提升网络运营效率。

8.5 中兴通讯 AIR Net 自智网络高阶演进解决方案

中兴通讯推出 AIR Net 自智网络高阶演进解决方案，从跨域协同、单域自治、内生智能三个层次出发，提供分层、分域、分级演进的自智能力，通过独立部署或者云部署，赋能网络数字化运营。

图 37 中兴 AIR Net 技术架构图

AIR Net 自智网络解决方案依托数智引擎和自智服务支撑智能化应用的设计、开发和部署。

自智服务负责汇聚各个领域的自智网络业务服务能力，比如质量优化服务、故障处理服务、变更监控服务、性能优化服务等，支撑自智网络应用层使用。

数智引擎主要由数据引擎、AI 大模型引擎、数字孪生引擎三大技术引擎构成，三者彼此相互协作支撑：

- 数据引擎为 AI 大模型引擎和数字孪生引擎提供所需的数据资源，并确保数据治理的有效进行。

- AI 大模型引擎通过大模型训推工具链、大模型库和智能体引擎，在模型训练、复用、智能体开发等多方面发挥着重要作用，构建自智网络的智能中枢。

- 数字孪生引擎通过多个内置子引擎的协同作用，实现对物理实体的全面映射、分析、优化管理。

这种相互赋能的关系使得三大技术引擎紧密协同工作，共同推进技术创新和发展。简而言之，数据引擎提供核心资源；AI 大模型引擎是网络智能中枢，对资源进行智能化处理；数字孪生引擎则将处理后的信息应用于实际场景中，形成一个闭环的、高效的创新生态。

8.6 京东云 JoyOps 智能运维

为了保障京东大促活动的正常进行，需要对业务系统各项性能和可用性指标进行全链路实时监控，当业务发生错误或者性能遇到瓶颈等问题时，需要能够迅速发现并定位根因，提升运维排障效率。京东 JoyOps 智能运维通过接入 AI 大模型能力，提供从移动 App、网页 H5 应用、小程序，到网关、后端服务和中间件的全链路监控服务，通过将专家语料库和 AI 大模型算法融合生成式故障诊断方案，在复杂的业务架构下也能实时掌握全栈性能情况，实现 1 分钟发现，5 分钟定位，10 分钟解决，提升服务稳定性。

图 38 京东云 JoyOps 智能运维示意图

JoyOps 智能运维构建了四大优势能力：

- AI 自动化运维：JoyOps 能够覆盖超过 90%的故障场景，并且准确率达到 80%以上，大部分常见故障可以自动检测并修复，极大减少了人工干预，提升运维效率。

- 故障自动预判：通过因果推理增强，JoyOps 可以预测并预防潜在的故障，故障响应时间显著缩短，提高业务的稳定性和连续性。

- 运维知识库沉淀：通过持续学习私域知识和和公开运维知识，JoyOps 不断优化自身，并将多模态运维数据整合，实现了跨部门和系统的数据一致性，使得运维人员员可以快速掌握和应用最佳实践。

- 全场景故障解决：提供无阈值监控、日志分析、根因分析、智能巡检、运维运营 AI 问答、慢 SQL 诊断与优化、硬件故障预测等运维场景能力。

目前，JoyOps 智能运维已大规模应用于京东内部场景，支持 618、11.11 等高并发和复杂业务场景，QPS 达千万级流量时，可用率仍可高达 99.99%，为业务应用稳定运行提供保障。

第 9 章　AI for Network 的挑战与未来趋势

本章旨在前瞻性地探讨 AI for Network 技术在演进过程中所面临的挑战，并深入分析其未来的发展趋势，为网络产业的持续创新与升级提供战略性思考与展望。

9.1 未来发展趋势

AI for Network 的未来发展趋势将围绕智能化、自主化、服务化和生态化展开，具体表现为大模型在网络领域的深度应用、自主进化的网络智能体系、网络即服务（NaaS，Network as a Service）模式的普及以及生态系统开放与协同创新。这些趋势将推动网络向更高效和更智能的方向演进，为全球数字化转型提供坚实支撑。

（一）大模型在网络领域的深度应用

随着 LLM 和多模态大模型技术的突破性进展，其在网络领域的应用将从辅助性工具向决策引擎演进。在网络运维方面，大模型将实现更高效的故障诊断与预测。通过对实时采集的网络数据进行深度分析，大模型能够精准识别异常模式，提前预警潜在故障，并快速定位故障根源，甚至提供详细的修复建议。此外，大模型在网络资源优化配置方面也展现出巨大潜力，它们能够评估不同配置组合对网络延迟、带宽利用率、可靠性等性能指标的影响，并通过模拟和优化算法找到最佳的网络配置方案，从而减少人工干预，提高网络配置的灵活性和适应性。随着多模态大模型的发展，未来网络大模型将能够融合文本、图像、语音等多种数据，实现更全面的网络状态感知和更智能的决策。

（二）自主进化的网络智能体系

未来的网络将不再是静态的、需要人工频繁干预的系统，而是具备自主进化能力的智能体系，这意味着网络智能体系将具备自我学习、自我适应、自我修复

和自我优化的能力，无需人工干预即可根据环境变化和业务需求进行持续演进。这种自主进化将体现在多个层面：在感知环节，网络智能体能够主动探索和发现新的数据源，提升对网络状态的全面感知能力；在分析和决策环节，AI 大模型能够根据新的数据模式和业务目标，自动调整和优化其算法，提升预测和决策的准确性；在执行环节，网络将具备高度自治的操作执行能力，在接收到决策指令后，能够自动触发修复流程、资源调度或配置更新，实现零接触式快速恢复的业务体验。

（三）网络即服务模式普及

网络即服务是一种将网络能力作为服务通过云端交付的模式，用户无需拥有、构建或维护自己的网络基础设施，而是按需订阅和使用网络功能。AI 的深度融合将加速 NaaS 模式的普及，并使其服务能力更加智能化、弹性化和个性化，能够根据用户的业务需求和应用场景，动态调整网络资源和策略，成为未来网络服务交付的主流范式。NaaS 的普及将推动网络从“拥有”向“使用”转变，使得网络能力像水电一样，成为按需取用的基础设施，从而加速各行各业的数字化转型。

（四）生态系统开放与协同创新

生态系统开放与协同创新将成为网络智能化升级的重要推动力。未来的趋势将是打破传统网络领域的封闭性，促进跨行业、跨领域的深度合作与创新。这包括提供标准化的接口和开发工具，吸引更多的第三方开发者、服务提供商和企业参与到网络能力的创新中来，共同构建丰富的应用生态。加速产学研深度融合，鼓励科研机构、高校与产业界紧密合作，共同攻克 AI for Network 领域的关键技术难题，加速科研成果向实际应用的转化。

9.2 战略建议与展望

面对 AI for Network 广阔的发展前景和伴随而来的挑战，需要从技术研发和产业发展等多个层面进行战略性布局和协同推进，以确保其健康、可持续发展。

（一）对技术研发的建议

（1）深化基础理论研究：加强对网络智能体的自主学习、多模态数据融合、因果推理、可解释 AI 以及分布式 AI 协同等基础理论的研究。这些核心理论的系统性突破与扎实推进，将构建起网络智能化的“理论护城河”，为技术创新提供源头活水。

（2）推动核心技术攻关：聚焦于数字孪生网络建模的精度与实时性、零接触故障自愈的准确性与效率、以及意图驱动网络的闭环控制与验证等关键技术。鼓励跨学科、跨领域的交叉研究，促进不同领域技术理念与方法的融合碰撞，加速关键技术突破与成果转化。

（3）构建开放创新平台：建立开放的创新平台和开源社区，提供丰富的网络数据集、标准化的 API 接口和易用的开发工具，吸引全球范围内的科研机构、高校和企业共同参与技术创新。通过开源等创新模式，加速技术迭代和应用孵化，形成良性循环的创新生态。

（二）对产业发展的建议

（1）加速应用场景落地：鼓励网络运营商、设备商和垂直行业企业紧密合作，推广 AI for Network 在实际网络运维、服务保障、安全防护等领域的典型应用场景。通过试点示范项目，验证技术成熟度，积累成功经验，形成可复制、可推广的解决方案。

（2）培育专业人才队伍：加大对 AI 网络复合型人才的培养力度，包括既懂网络又懂 AI 的工程师、架构师和研究人员。通过校企合作、产教融合等方式，建立多层次、多类型的培训体系，为产业发展提供坚实的人才支撑。

（3）构建产业生态联盟：推动产业链上下游企业、科研机构、行业组织等共同组建 AI for Network 产业联盟，协同制定技术标准、共享最佳实践、联合开展市场推广。通过构建开放、合作、共赢的产业生态，共同应对挑战，拓展市场空间。

第三部分 未来展望

第 10 章 AI 网络发展十大趋势

AI 网络技术正迎来从技术协同到生态共建的关键发展阶段。Network for AI 与 AI for Network 两大方向深度融合，在技术层面形成闭环支撑，在应用层面实现创新突破，在产业生态层面助力全局协同，其演进趋势将重塑智能时代的网络形态，为人工智能规模化发展与数字经济升级奠定基础。

AI 网络作为支撑人工智能规模化落地与数字经济深化发展的新型基础设施核心支柱，正在从“支撑 AI 需求”向“定义 AI 未来”跨越。立足当下技术发展与场景变革，AI 网络技术将呈现以下十大趋势，共同塑造智能时代的网络新形态。

(一) 从通用互联到智算中心网络范式

AI 大模型训练与超大规模推理对网络提出极致性能需求，推动网络从传统数据中心的“通用连接架构”向“智算中心专用网络范式”升级。通过物理拓扑、传输协议与通信库的协同设计，实现端到端性能优化，构建超高吞吐、超低时延、线性扩展的“神经网络骨架”，支撑千卡至万卡级智算集群的高效协同，成为 AI 算力释放的核心基座。

(二) 从独立层级到超融合无中心架构

传统云边端层级化架构难以适配分布式 AI 的实时协同需求，将演进为超融合无中心架构。基于 Mesh/Torus 多维互联与智能路由，打破计算、存储、网络资源的层级壁垒，实现全域资源的按需调度与任意节点高效通信，为边缘 AI 协同、分布式训练提供韧性更强的底层支撑。

(三) 从尽力而为到确定性智能协议栈

传统 TCP/IP 的“尽力而为”特性与工业控制、自动驾驶等 AI 场景的确定性需求存在根本矛盾，驱动协议栈向“AI 感知的确定性智能体系”演进。融合 TSN/DetNet 时间敏感技术、意图感知能力，在网络层、传输层构建可预期、可测量、可保障的行为范式，为关键 AI 业务提供实时 SLA 保障。

(四) 从云为核心到泛在云边端智能协同

AI 应用正从云端集中式部署向物理世界末梢渗透，推动网络向“云边端智能协同架构”转型。依托 5G/6G、卫星互联网与 LPWAN 泛在连接，实现智能任务在云、边、端之间的动态分解、模型迁移与推理结果融合，在保障数据隐私与实时性的同时，最大化分布式算力价值，赋能智能制造、智慧城市等场景的深度智能化。

(五) 从被动配置到意图驱动智能调度

静态网络配置无法适配 AI 负载的动态特性，将被“意图驱动智能调度”取代。基于强化学习与全局感知，网络可实时解析业务 SLA 需求与算力、带宽、拓扑等资源状态，实现毫秒级资源编排、流量调度与路径优化，完成从“人适配网”到“网适配 AI”的范式转变，支撑 AI 工作流的弹性伸缩。

(六) 从数据集中到隐私优先联邦协同范式

数据隐私法规与数据孤岛问题倒逼 AI 训练推理模式革新，推动网络支撑隐私优先的联邦协同范式。通过加密参数传输、跨域安全隔离等技术，实现客户端、边缘、多云等分布式节点的模型协同训练推理，在不泄露原始数据的前提下打破数据壁垒，成为医疗、金融等敏感领域 AI 落地的核心使能技术。

(七) 从功能附加到原生智能网络内核

网络智能化将从外挂式分析引擎向原生智能内核演进。轻量化 AI 模型与推理引擎将深度嵌入交换机、网卡等设备的数据面，实现流量特征实时感知、异常行为本地决策、策略自主执行，使网络具备分布式自优化能力，为实现自驱动网络奠定基础，大幅提升自动化与故障响应效率。

(八) 从千网一面到知识内化行业专网

通用网络服务难以满足垂直行业的差异化 AI 需求，将演进为“知识内化的行业专网”。网络深度融合工业控制逻辑、医疗隐私规范、车联网移动性特征等垂直行业知识，定制化网络架构、协议与管理策略。例如，工业专网通过确定性传输保障 AI 控制指令实时响应，医疗专网以强化隐私保护支撑 AI 安全应用于诊疗，车联网专网凭超低时延赋能自动驾驶 AI 毫秒级决策，其定制化能力让 AI 在各领域价值最大化。

(九) 从传统互联网到智能体互联网

传统互联网以人类为通信主体，实现全球信息互联。随着通用人工智能技术的突破性进展，具备自主决策、环境感知与交互能力的智能体正成为网络新主体，推动互联网向智能体互联网（IoA，Internet of Agents）跃迁。通信维度将从人与人互联扩展至智能体与智能体互联以及人与智能体互联，这不仅大幅拓宽网络的内涵与边界，更将引发产品服务、产业生态乃至社会形态的深刻变革，开启“万物皆智能体”的新纪元。

(十) 从算网分离到算网电融合共生

AI 算力的指数级增长与双碳目标共同推动算网电融合共生。网络作为枢纽，将协同算力布局与电力供应（尤其是绿电），通过全局能效优化算法动态调度任务至绿电富集区，同时推动自身向光电混合、液冷背板等高能效方向演进，支撑“东数西算”等国家级工程落地，实现 AI 算力的绿色可持续发展。

参考文献

[1]. IDC，2025 年中国人工智能计算力发展评估报告

[2]. 未来网络白皮书，智算网络技术与产业白皮书

[3]. 百度智能云，智算中心网络架构白皮书

[4]. 中国移动通信研究院，全向智感互联 OISA 技术白皮书

[5]. 中国移动通信研究院，全调度以太网技术架构白皮书

[6]. IMT-2030（6G）推进组，6G 典型场景和关键能力白皮书

[7]. 王光全，满祥锟，徐博华，等. 确定性光传输支撑广域长距算力互联[J]. 邮电设计技术，2024(2): 7-13.

[8]. 华为，确定性 IP 网络介绍

[9]. 中兴，超节点技术：NVL72 和 ETH-X

[10]. Ultra Accelerator Link Consortium, Inc. (UALink). (2025). UALink_200 Rev 1.0 Specification

[11]. Ultra Ethernet Consortium. (2025). Ultra Ethernet specification: v1.0.

[12]. Broadcom Corporation. (2025). Scale Up Ethernet Framework Specification (Scale-Ethernet-RM102).

[13]. Zuo P, Lin H, Deng J, et al. Serving Large Language Models on Huawei CloudMatrix384 [J/OL]. 2025-06-15. https://doi.org/10.48550/arXiv.2506.12708.

[14]. Huang Y, Huang T, Zhang X, et al. CSQF-based Time-Sensitive Flow Scheduling in Long-distance Industrial IoT Networks [J/OL]. 2024-09-15. https://doi.org/10.48550/arXiv.2409.09585.

[15]. Qian K, Xi Y, Cao J, et al. Alibaba HPN: A Data Center Network for Large Language Model Training [C]//Proceedings of the ACM SIGCOMM 2024 Conference. Sydney, NSW, Australia: ACM, 2024: 1-16. DOI: 10.1145/3651890.3672265.

[16]. TM Forum, 自智网络产业白皮书 6.0

[17]. 亚信科技, AI Agent 赋能自智网络技术探析与实践

[18]. 未来移动通信论坛, 10.0I 意图驱动自智网络

[19]. 宋延博, 高先明, 杨春刚, 等. 意图驱动的韧性网络安全研究[J]. 系统工程与电子技术, 2024, 46(9):3211-3220

[20]. 李福亮, 范广宇, 王兴伟, 等. 基于意图的网络(IBN)研究综述[J]. 软件学报, 2020. DOI: 10.13328/j.cnki.jos.006088.

[21]. Pang, L., Yang, C., Chen, D., Song, Y., & Guizani, M. (2020). A survey on intent-driven networks. IEEE Access, 8, 22862-22873. doi: 10.1109/ACCESS.2020.2969208

[22]. 孙滔, 周钺, 段晓东, 等. 数字孪生网络(DTN):概念, 架构及关键技术[J]. 自动化学报, 2021, 47(3):14.DOI:10.16383/j.aas.c210097.

[23]. 中国移动研究院, 基于数字孪生网络的 6G 无线网络自治白皮书

[24]. 中兴, 2025 中兴通讯自智网络白皮书

[25]. 中金, AI 十年展望: AI Agent 元年已至, 应用拐点或将到来

[26]. 段晓东, 孙滔, 陆璐, 等. 智能体互联网: 概念、架构及关键技术[J/OL]. 电信科学.https://link.cnki.net/urlid/11.2103.TN.20250715.1430.002

[27]. 陈天骄, 刘江, 黄韬. 人工智能在网络编排系统中的应用[J]. 电信科学, 2019 (5). DOI: 10.11959/j.issn.1000-0801.2019095.

[28]. 潘彤, 余文艳. 智能体互联网：未来网络的新图景[J]. 中国电信业, 2025(4): 30-34.

缩略语

中文全称	英文全称	英文缩写
人工智能	Artificial Intelligence	AI
大语言模型	Large Language Model Technology	LLM
确定性网络	Deterministic Network	DetNet
意图驱动网络	Intent-Driven Network	IDN
数字孪生网络	Digital Twin Network	DTN
深度学习	Deep Learning	DL
分布式训练网络	Distributed Training Network	DTN
软件定义网络	Software Defined Networking	SDN
网络功能虚拟化	Network Functions Virtualization	NFV
计算机视觉	Computer Vision	CV
自然语言处理	Natural Language Processing	NLP
图形处理器	Graphics Processing Unit	GPU
传输控制协议	Transmission Control Protocol	TCP
网际协议	Internet Protocol	IP
远程直接内存访问	Remote Direct Memory Access	RDMA
神经网络架构	Transformer	TF
机器学习	Machine Learning	ML
超以太网联盟	Ultra Ethernet Alliance	UEC
可靠无序	Reliable Unordered Delivery	RUD
可靠有序	Reliable Ordered Delivery	ROD
可靠无序幂等	Reliable Unordered Delivery For Idempotent	RUDI

中文全称	英文全称	英文缩写
不可靠无序	Unreliable Unordered Delivery	UUD
互联网工程工作小组	Internet Engineering Task Force	IETF
确定性 IP 网络	Deterministic IP	DIP
指定周期排队转发	Cycle Specified Queuing and Forwarding	CSQF
循环排队转发	Cyclic Queuing and Forwarding	CQF
时间敏感网络	Time-Sensitive Networking	TSN
光传送网络	Optical Transport Network	OTN
光交叉连接	Optical Cross-Connect	OXC
高带宽域	High Bandwidth Domain	HBD
时空图神经网络	Spatio-Temporal Graph Neural Network	STGNN
深度神经网络	Deep Neural Networks	DNN
安全编排和自动化响应	Security Orchestration, Automation and Response	SOAR
联邦学习	Federated Learning	FL
网络即服务	Network as a Service	NaaS
智能体互联网	Internet of Agents	IoA