

智算中心网络技术发展与应用



Evolution and Applications of Network Technology in Intelligent Computing Center

段威/DUAN Wei^{1,2}, 李和松/LI Hesong^{1,2}, 周昆/ZHOU Kun¹

(1. 中兴通讯股份有限公司, 中国 深圳 518057;
2. 移动网络和移动多媒体技术全国重点实验室, 中国 深圳 518055)
(1. ZTE Corporation, Shenzhen 518057, China;
2. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China)

DOI: 10.12142/ZTETJ.202406007

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250109.1000.006.html>

网络出版日期: 2025-01-09

收稿日期: 2024-10-16

摘要: 从应用子层、网卡子层、网络子层以及管控子层构成的完整技术栈出发, 介绍了智算中心网络的关键技术。在分析智算中心网络发展趋势的基础上, 介绍了中兴通讯在坚持核心自研的原则下, 在芯片、产品和组网方案等方面开展的一系列创新。认为面向人工智能 (AI) 场景优化将成为智算中心网络发展的关键因素, 行业必须在基础芯片、设备形态、网络架构、网络协议以及应用生态等方面做出更多努力, 进一步推进算侧、端侧和网络侧关键技术的融合发展。

关键词: 智算中心网络; 大模型; 以太网

Abstract: The key technologies of the intelligent computing center network are introduced from four aspects: application sublayer, network card sublayer, network sublayer, and control sublayer. ZTE Corporation has carried out a series of innovations in chip, product, and networking solutions while adhering to the principle of autonomous research and development. It is believed that artificial intelligence (AI) scenario optimization will become a key factor in the development of intelligent computing center networks, and the industry must make more efforts in the basic chip, device form, network architecture, network protocols, and application ecology to further promote the integration and development of key technologies on the computing side, end side, and network side.

Keywords: intelligent computing center network; large language model; ethernet

引用格式: 段威, 李和松, 周昆. 智算中心网络技术发展与应用 [J]. 中兴通讯技术, 2024, 30(6): 39-47. DOI: 10.12142/ZTETJ.202406007

Citation: DUAN W, LI S H, ZHOU K. Evolution and applications of network technology in intelligent computing center [J]. ZTE technology journal, 2024, 30(6): 39-47. DOI: 10.12142/ZTETJ.202406007

以大语言模型 (LLM) 为基础的生成式人工智能 (AI) 技术因其良好的通用性与泛化能力, 正在快速引领新一轮的科技革命和社会产业的变革。当前越来越多的科技公司竞相推出千亿、万亿参数规模的通用和垂直大模型^[1], 例如 OpenAI 的 GPT 系列、Meta 的 LLaMA 系列、百度的文心一言大模型、阿里巴巴的千问大模型等。当模型的参数规模超过数百亿后, AI 大模型的语言理解能力、逻辑推理能力以及问题分析等能力迅速提升。以 GPT3.5 为例, 参数规模达 1 750 亿, GPT4 的参数规模更是达到了 1.8 万亿, 参数规模的增加提升了大模型处理复杂问题的能力。训练这类超大参数规模的大模型给智能计算基础设施带来了前所未有的挑战, 通常需要几千甚至数万张图形处理器 (GPU) 加速卡并行协同工作^[2]。基于数据并行、流水线并行和张量并行等多种并行技术的分布式并行计算是实现 AI 大模型训练的关键

手段。业界普遍认为支撑万卡以上规模的 GPU 集群高效运行的瓶颈在于网络互联能力^[3]。如何提供一种无损、超高带宽、超低延迟、超高稳定性且可高度扩展的网络互联方案逐渐成为行业研究和关注的焦点。

1 智算中心网络特点和性能需求

大量的理论分析和工程实践表明, 为确保昂贵的算力集群资源的高效利用, 面向 GPU 互联的智算中心网络在组网架构、流量模型和性能指标 3 个方面和传统的数据中心网络有所不同。因此, 我们需要推动智算中心网络性能的跨越式发展。

1.1 智算中心网络的组网架构特点

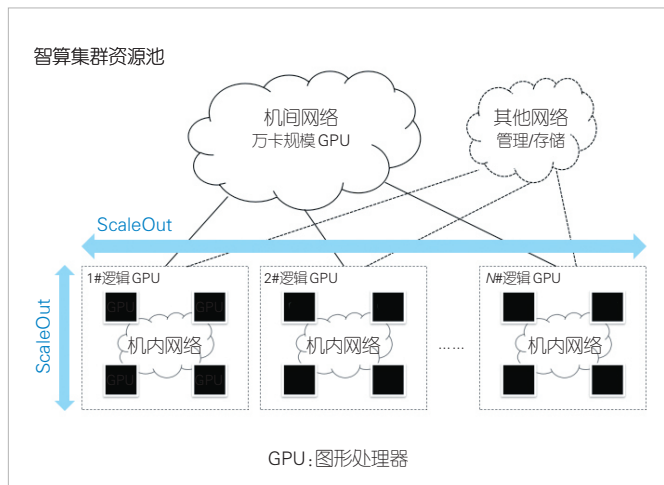
如图 1 和表 1 所示, 与传统数据中心不同的是, 除了传

统的管理、存储等通用网络以外，智算集群通常存在机内网络和机间网络。其中，机内网络实现的是有限数量（通常为几个到几百个）GPU之间超高速互联，通常称为ScaleUp网络。其主要目的是将多个GPU组织成一个逻辑上的超级GPU，以满足大模型并行切分的需求。此类网络通常需要数百GB的超高通信带宽、几百纳秒的低延迟以及一定的缓存一致性要求。典型的例子是英伟达通过NVlink/NVswitch构建256卡的SuperPOD。机间网络的目的是构建超大规模的GPU集群（如万张GPU卡集群规模），通常被称为ScaleOut网络。该类网络通常需要单端口100 Gbit/s以上的无收敛组网技术，由网卡和交换机组成。典型的此类网络为InfiniBand（IB，一种用于高性能计算的计算机网络通信标准）网络或者基于融合以太网的远程直接内存访问（RoCE）网络。

在实际业务运行的过程中，机内网络和机间网络由一套高性能通信库来实现灵活的资源调配。由此可见，智算中心是网络一个高度复杂的软硬件技术栈，而不是一个单点技术或者产品。

1.2 智算中心网络的流量模型特点

智算集群的业务特性决定了其在网络侧呈现独特的流量



▲图1 智算中心网络的组网架构

▼表1 智算网络的分类和定位

网络类型	网络定位	网络特点	典型技术
机内网络	ScaleUp网络,将多个GPU组成一个逻辑上的超级GPU	超高带宽,GPU点到点通信带宽百GB以上;扩展性受限,通常为几百个GPU以内;超低延迟;缓存一致性等	NVlink/Nvswitch
机间网络	ScaleOut网络,实现集群的横向扩展,组成超大的智算集群	单端口超过100 Gbit/s无收敛组网,由网卡和交换机组成,通常要求支持RDMA	Infiniband、Ethernet
其他网络	管理/存储/用户南北向流量	有收敛比的通用数据中心网络	Ethernet

GPU:图形处理器 RDMA:远程直接内存访问

模型：低熵、大象流、同步效应。其中，低熵表示同一时刻网络中的活跃数据流（通常由五元组决定）数量较少，通常为千级，远远低于通算集群动辄数万甚至数十万的流数，这将导致传统基于流的负载均衡技术在智算中心网络中不再适用；大模型的参数规模和GPU的超高并发处理能力导致智算中心网络中大象流占据主导地位，根据中兴通讯的现网统计，单条流的峰值带宽即可达到上百GB；同步效应指的是AI训练和推理任务充斥着大量的集合通信，同一任务内的不同流之间存在明显的同步需求，导致网络的长尾延迟对业务性能影响被显著放大^[5]。

智算中心网络这种独特的流量模型会直接影响到现有技术的有效性，这也是推动智算中心网络技术创新的原动力。

1.3 智算中心网络的性能指标需求

结合AI大模型的业务特点和流量模型特征，中兴通讯从组网规模、带宽、延迟、可靠性以及可管理性等几个维度对智算中心网络的性能指标需求开展定性和定量分析。

1) 性能指标定性分析

LLM训练、LLM推理、智能推荐等常见AI业务场景对网络的定性需求如表2所示，其中带宽指的不仅仅是网络物理带宽，还指基于物理网络能达到的业务有效吞吐；时延指的是业务转发延迟，包括网络静态转发延迟以及由于排队、丢包重传导致的动态延迟；尾延迟是评价网络延迟性能稳定性的重要指标，在AI场景中对业务性能影响巨大。

2) 性能指标定量分析

从网络的视角来看，我们通常希望同样的产品和解决方案能够尽可能覆盖多样化的业务场景。综合考虑业务场景的需求，中兴通讯提炼出5个方面的定量指标，同时也分析出

▼表2 典型人工智能业务场景网络需求

业务场景	集群规模	带宽	时延	尾延迟	稳定性	可管理性
LLM训练	超大	超高	一般	较高	超高	超高
LLM推理	较小	一般	较高	较高	高	高
智能推荐	较小	较高	一般	一般	高	高

LLM:大语言模型

了当前网络实际能力与这些指标的差距，并给出了优化方向，具体如表3所示。

2 智算中心网络的关键技术和发展趋势

智算中心网络是一个高度复杂的软硬件技术栈。本章中，我们将从关键技术和行业趋势两个方面展开具体分析。

2.1 智算中心网络的关键技术

如图2所示，从技术架构层面分析，智算中心网络主要包含应用、网卡、网络和管控4个技术子层。为了满足AI业务的性能需求，各个子层需要引入如下一些核心技术。

1) 应用子层核心技术

在智算中心网络的应用场景中，网络能力通常被封装成高性能通信库供上层应用逻辑调用，典型的通信库如lib-fabric、消息传递接口(MPI)、英伟达集合通信库(NCCL)等。通信库通常需要根据具体的物理网络能力进行适配才能达到最佳性能，如拓扑发现、拓扑亲和性、负载均衡路径发现、集合通信加速等，这是智算中心网络不可或缺的一个环节。

2) 网卡子层核心技术

网卡是智算中心网络流量的起点和终点，很多关键功能都必须依托网卡来构建。网卡子层主要包括如下关键技术。

(1) RDMA协议及其优化

远程直接内存访问(RDMA)是网络实现高吞吐性能的基础。然而，在大规模组网环境下，传统RDMA协议容易成为性能瓶颈^[6]。为了在大规模集群网络中实现高性能的RDMA传输，需要对网卡进行一系列优化，包括协议设计和传输模式。这些优化旨在解决传统RDMA协议在扩展性方面的不足。例如北京邮电大学提出的FedRDMA^[7]，将模型分块之后再使用RDMA传输，优化了跨数据中心联邦学习场景中的通信效率；南京大学提出了一种可扩展的RDMA传输方法^[8]，通过有效共享连接来提高性能。

(2) 新型拥塞控制算法



▲图2 智算中心网络的关键技术

拥塞控制是高性能网络要解决的重要技术问题。传统拥塞控制算法如数据中心量化拥塞通知(DCQCN)^[9]在扩展性、流公平性以及时延性能方面存在瓶颈，需要新型拥塞控制算法的创新。例如，阿里巴巴提出的高精度拥塞控制(HPCC)^[10]，通过利用In-band网络遥测，实现了高精度的拥塞控制，显著提升了网络的吞吐量并降低了延迟性能。Google提出的Swift^[11]拥塞控制算法，依靠基于延迟的反馈机制，进一步降低了延迟，并提高了吞吐量。Poseidon^[12]通过使用带内遥测(INT)携带拥塞信息，只响应网络中瓶颈点的拥塞信号，从而达到了高效拥塞控制的目的，在快速收敛的同时实现最大公平性。新加坡国立大学与诺基亚合作提出的LinkGuardian^[13]则采用链路本地重传机制，有效解决了数据中心网络中的链路丢包问题，减少了丢包率，并维持了较低的延迟和高吞吐量。中国科学技术大学提出的低时延高精度拥塞控制(LHCC)^[14]使用带外遥测技术，能够快速通知网络状态，使发送端在一个往返时延(RTT)内感知到拥塞情况，并根据路径上所有队列的状态调整发送速率。这种机制使得LHCC在网络中存在多个瓶颈时依然能提供更精确的拥塞控制。中国科学院计算技术研究所提出的FlexPass^[15]是一种基于Credit的传输协议，结合加权公平队列和双控制循环，在保持低延迟的同时实现高吞吐量和零丢包。这些新型拥塞控制算法通过引入高精度监控和智能反馈机制，进一步推动了数据中心网络的性能提升，特别是在减少延迟、提高

▼表3 智算网络定量需求分析

智算网络需求	需求描述	当前数据中心网络能力	优化方向
万卡集群规模	按照GPU节点数计算,集群规模按照场景分级,私有云资源千卡级,AI工厂/公有云万卡级	传统RDMA集群数小于1000	智能网卡、物理网络
TB互联带宽	单台8卡服务器接入带宽1.6T+,有效吞吐>90%	接入10G/25G,有效吞吐<60%	物理网络、负载均衡
极致网络性能	0丢包,微秒级时延,微秒级低抖动	时延亚毫秒级,拥塞情况下秒级,基础无损	网络拓扑、拥塞控制
超高网络稳定性	亚毫秒级故障恢复,性能一致性	50ms级检测,秒级/分钟级收敛,性能不稳定	管控、网络设备
多维网络自动化	部署、验收、运维、变更自动化	能力参差不齐,整体欠缺	管控、网络设备

AI:人工智能 GPU:图形处理器 RDMA:远程直接内存访问

吞吐量和减少丢包方面取得了显著进展。

(3) 网络多路径控制

网络多路径是智算中心网络吞吐性能的基础，也是在故障场景下保障通信性能的重要基础手段。路径发现、探测以及报文喷洒是网络多路径控制的主要功能。新加坡国立大学提出的 ConWeave^[16]通过网络内的重路由和数据包重新排序，为 RDMA 流量提供了有效的负载均衡，显著改善了平均和高百分位的流完成时间。湖南大学的 RDMA 轻量级快速报文重排机制 (LEFT)^[17]则利用双状态共享位图方案减少了内存消耗，并通过快速和慢速路径的数据包重新排序降低了延迟，即使在多路径 RTT 差异较大的情况下，仍能保持高吞吐量。复旦大学的基于主机的流片微调 (HF2T)^[18]进一步优化了 RDMA 负载均衡，通过在主机端延迟少量数据包，延长数据包间的时间间隔，增加 flowlet 的生成机会，从而提升 flowlet 负载均衡的效果。清华大学基于向量协议的 RDMA 拥塞感知负载均衡 (CAVER)^[19]则采用了基于向量协议的拥塞感知负载均衡方法，通过利用确认字符 (ACK) 包在网络中传播拥塞信息，实时为源服务器机柜顶交换机 (ToR) 提供最不拥塞的路径。在不改变现有硬件的前提下，CAVER 通过对 ACK 包携带的路径拥塞信息进行传播，帮助源交换机快速找到最优路径。

3) 网络子层核心技术

网络层是由交换机组成的高速互联网络。为了更好地满足智算业务的实际需求，网络本身需要引入一系列的技术创新，包括新型网络拓扑、新型网络协议、新型组网技术、新芯片能力、亚毫秒级故障自愈以及在网计算等。这些技术涉及网络芯片、网络设备、网络协议等多个方面，也是当前技术创新最活跃的领域之一。

(1) 新型网络拓扑

智算中心的网络拓扑有很多技术选择，包括 CLOS、FatTree、BiGraph^[20]、Dragonfly^[21]以及 Torus^[22]拓扑，每一种拓扑都有其各自的优点和适用场景。新型网络拓扑及其配套的路由协议是智算中心网络需要关注的重要内容。

(2) 新型网络协议

网络组网拓扑和性能需求的变化必然导致新型网络协议的出现。为了能在新场景中实现最佳的网络路由、高效的负载均衡，需要设计新的网络协议。这也是目前业界研究的重点，典型的协议如胖树网络路由协议 (Rift)^[23]、自适应路由^[24]等。

(3) 新型组网技术

针对智算中心网络的组网需求，目前业界提出了很多新型组网技术，包括分布式分散式机框 (DDC)^[25]、全调度以

太网 (GSE)^[26]等。他们大多通过一种封闭式 Fabric 设计以期达到最优的网络性能。该领域也是当前创新最为活跃的领域。

(4) 新芯片能力

网络芯片的优化空间广泛，除了传统的可编程能力外，Buffer 管理和时延优化是两个重要的方面。目前明确采用共享缓存架构的厂商有思科、博通以及英伟达，且作为其主要的架构亮点，低时延是需要持续优化改进的方向，其本质是采用更低层次的 Cut-Through 逻辑。

(5) 亚毫秒级故障自愈

为了提高 AI 算力集群的资源利用率，网络需要具备能够达到亚毫秒级的故障自愈能力。由于智算中心网络通常都有多条等价路径可用，收敛性能的瓶颈则在于故障感知能力。当前基于 echo-reply 机制的故障检测基本都是 50 ms 级，需要芯片层引入一些快速故障感知甚至故障预测的能力。

(6) 在网计算

在网计算是当前智算中心网络研究的热点，也是英伟达网络互联方案的核心。通过网络交换机支持集合通信卸载加速，可以获得计算任务提速、网络拥塞缓解的双重收益。根据英伟达给出的测试数据，其在网计算的实现方案可扩展的分层聚合与规约协议 (SHARP)^[27]可带来 2 倍以上的计算任务加速，其集群规模越大，收益越明显。在未来机内互联和机间互联的场景中，在网计算带来的整体收益将日益明显，这也将成为未来智算中心网络的关键技术需求之一。

4) 管控子层核心技术

对于上万节点规模的智算集群而言，网络的部署、测试、验收、运维和变更等需要依托管控平台构建完善的自动化能力。这种管控平台的管理对象不再限制为网络本身，而是逐渐覆盖到应用、网卡和网络等端到端的通信链条，这也是与传统网络的管控系统最大的区别。

各子层所涉及的关键技术的基本原理以及与需求的对应关系如表 4 所示。

2.2 智算中心网络的发展趋势

AI 大模型让智算产业空前繁荣，智算中心网络也进入了技术创新的高速发展周期。通过对当前行业和技术洞察分析，中兴通讯认为当前智算中心网络呈现如下几个明显的发展趋势。

趋势 1: RDMA 成为智算中心网络高性能协议的主流技术

传统数据中心业务通常对吞吐和延迟有较高的容忍度，简单易用的传输控制协议/互联网协议 (TCP/IP) 即可满足

▼表4 智算网络的关键技术分析

场景	需求	关键技术	功能主体	技术原理
智算网络	万卡级集群规模	新型网络拓扑	交换机	面向智算的特定场景(如LLM训练),在交换机能力一定的情况下,通过新型网络拓扑提升集群规模和网络效率
		RDMA 协议优化	网卡	针对传统 RDMA 协议的不足,开展协议优化,以支撑集群向万卡以上规模扩展(如从 RC 传输模式转向 RD 传输模式等)
	TB 级互联带宽	TB 级接入	交换机	基于大容量交换机芯片支持 100G/400G/800G 高密度组网
		网络多路径能力	网卡、交换机	通过端网协同完成流量的多路径传输,提升网络吞吐性能
	极致网络性能	新型拥塞控制算法	网卡	通过拥塞控制算法的创新,确保业务能在集群规模扩展的前提下低延迟和低抖动
		新芯片能力	交换机	在交换芯片层面针对 AI 场景引入新的特性,增强故障倒换和拥塞感知能力
		新网络协议	交换机	针对新型网络拓扑和新的拥塞故障场景,研究对应的网络协议
	超高稳定性	新型组网技术	交换机	研究新型组网技术如 DSF、全调度以太网 GSE 等
		在网计算	网卡、交换机	网卡和交换机协同设计,实现集合通信的卸载加速功能
			故障快速检测	交换机
		性能一致性	交换机	研究网络硬隔离和软隔离的技术方案

AI: 人工智能
DSF: 分布式调度网络
GSE: 全调度以太网
LLM: 大语言模型
RC: 可靠连接模式
RD: 可靠数据报模式
RDMA: 远程直接内存访问

业务的承载需求。智算中心网络本质上一个高性能网络的需求。传统 TCP/IP 协议已经无法满足业务对高吞吐、低延迟、零算力损耗的要求，RDMA 逐渐成为智算中心网络高性能协议的主流技术。智算中心网络未来发展的核心目标之一是支持超大规模的 RDMA 高性能集群组网。

趋势2：采用以太网来构建智算中心网络成为更广泛的行业共识

以太网已成为智算中心网络最主流的支撑技术。过去10年里，以太网凭借完善的标准体系、成熟的产业链、快速演进的接口速率、灵活的向后兼容能力等优点，成为当前数据中心网络采用的事实标准。目前以太网接口已经发展到800G/1.6T，交换芯片容量从100G迅速提升到51.2T/102.4T，容量增长近100倍，单比特功耗下降90%以上。将以太网延伸到智算中心网络各个场景已经成为行业共识，面向AI场景优化的新型以太网技术正处于快速发展期。在这种趋势的引领下，国际上成立了超级以太网联盟（UEC），中国也出现了全调度以太网和高通量以太网等一系列创新成果。相信以太网将成为未来智算中心网络最基础、最主流的支撑技术。

趋势3：场景融合成为智算中心网络技术的创新方向

随着AI集群组网规模的持续增长，多场景融合将是未来智算业务对网络的内在需求。场景融合具体体现在如下几个方面。

1) 总线和网络的融合

当前AI集群网络由负责垂直扩展的总线型互联和负责横向扩展的网络互联构成两个独立异构的互联域。当集群规模进一步扩展到十万节点甚至百万节点的规模时，这种组网

方式的成本和可维护性将难以为继。随着互联技术的快速演进，总线和网络的统一承载成为可能，网络总线化和总线网络化的趋势将成为智算中心网络技术创新的主要方向。

2) 多业务的融合承载

智算业务的发展日新月异，混合专家模型（MoE）和多模态技术会带来不同的网络侧需求，流量模型、带宽需求、时延指标等呈现多样化的特征，需要网络具备“一网多用”的综合承载能力。多业务综合承载不仅是未来智算中心网络面临的挑战，也是需要技术创新解决的核心问题之一。

3) 光电技术融合组网

从英伟达最新的NVL72/576互联架构^[28]来看，电互联和光互联融合组网是构建高能效比超级算力集群的关键。光电混合组网技术在以谷歌为代表的下一代数据中心大规模部署落地^[29]。大量的实践数据表明，光电混合组网在构建高性能、低能耗、低成本的AI网络中效果明显^[30]，是下一代智算中心网络架构演进的趋势。

场景融合必然带来网络技术的跨域融合，这也是是智算中心网络下一阶段创新的热点方向。

趋势4：新型大容量网络芯片成为智算中心网络发展的基石

大容量网络芯片是AI数据中心高速互联的基础载体，主要包括数据处理器（DPU）网卡芯片和网络交换芯片。DPU网卡芯片是流量的网络入口，是RDMA、拥塞控制、各种加速引擎的功能载体。网络交换芯片数据中心交换机最核心的部件，决定着网络的组网规模和整体性能。当前国际最先进的DPU网卡芯片已经具备400G/800G接口速率，51.2T的交换芯片已经规模化落地，可以支撑3万张新一代GPU集

群。随着智算业务对高速互联的需求持续攀升，网络芯片正处于一个高速发展的阶段，呈现出如下明显的发展趋势。

1) 容量持续增长，单比特功耗持续降低

过去 10 年，以太网交换芯片容量从百 G 迅速提升到 51.2T，容量增长近 100 倍，单比特功耗下降 90% 以上。在 AI 的驱动下，未来交换网络芯片容量将迅速突破 100T，单比特功耗进一步降低。与此同时，400G/800G DPU 网卡需求也将迎来井喷。

2) 面向 AI 场景优化成为网络芯片发展的基本要求

在过去两年里，新一代网络芯片引入面向 AI 场景优化的新特性成为行业主旋律。这也将是未来 5~10 年网络芯片更新迭代的主要推动力。典型技术包括超低延迟、故障预测、智能流分析引擎、基于容器/包的负载均衡、在网计算等。

面向 AI 场景优化的新型大容量网络芯片是智算中心网络发展的基石，需要在高速接口、交换架构等基础技术方面持续创新突破，是自主可控需要重点关注的核心技术之一。

3 智算中心网络的典型应用场景和解决方案

智算中心网络在应用场景上存在一定的多样性，行业解决方案也呈现百家争鸣的繁荣景象。中兴通讯做了大量的实践和理论分析，本章节我们将对应用场景和行业解决方案分别进行概述。

3.1 智算中心网络的典型应用场景

按照业务类型分，除了传统通算场景已经覆盖的虚拟私有云（VPC）网络、存储网络和管理网络以外，智算中心网络聚焦在 AI 加速器（包括 GPU 和 DPU 等）之间的高速互联网络，包括通常意义上的机内 ScaleUP 网络和机间 ScaleOut 网络两部分。中兴通讯认为，应用场景需要分类分级，准确

把握应用场景的需求是匹配最佳网络方案的基础，具体如下。

1) 不同业务类型对网络的需求差异较大

如表 2 所示，典型的业务类型包括 LLM 训练、LLM 推理、智能推荐等，不同的业务对网络性能的要求不同^[1]。具体而言，LLM 训练是典型的带宽密集型业务，延迟容忍度较高，流量的轨道效应明显；LLM 推理的 Decode 是延迟敏感性业务，延迟通常决定着推理集群的资源利用率；智能推荐通常要求较高的带宽和延迟性能。这种业务类型的差异会直接影响网络方案的选择。

2) 用户的目标集群规模分级明显

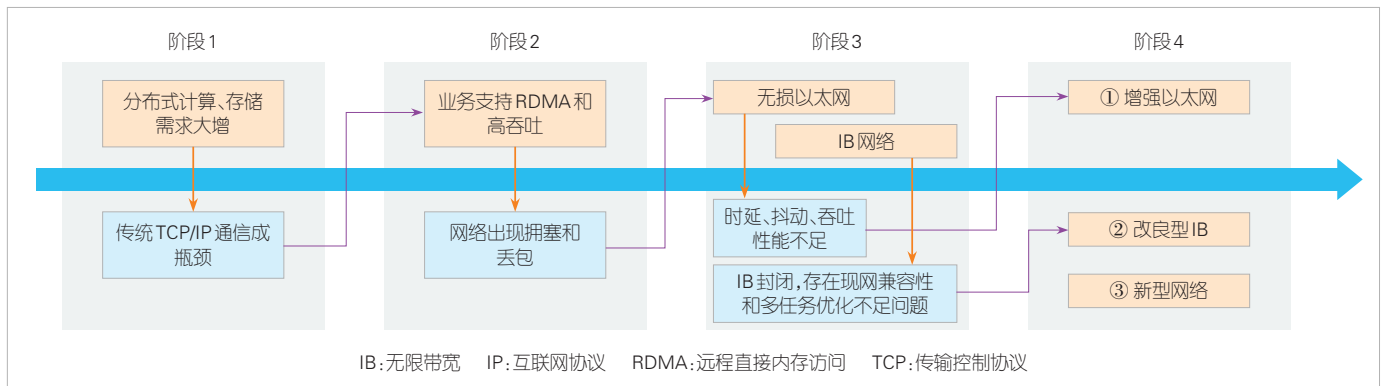
当前行业掀起了万卡以上甚至十万卡、百万卡规模智算集群研究的热潮，而真正需要万卡以上规模集群的应用通常是 LLM 基础训练。据统计，未来真正有能力投资建设万卡以上资源池来从事基础大模型训练的企业不到 10%。这意味着 90% 以上的企业建设智算资源池是用来进行大模型的精调和推理，而满足这类需求的集群规模通常只需要千卡级以下规模的集群。在千卡以下的智算集群中，网络方案可以做得极其简单，过于复杂的方案会导致用户付出额外成本。中兴通讯认为，未来智算中心网络的建设将进入一段时间的冷静期，从用户的实际需求出发匹配最具性价比的方案才是网络价值的回归。

3.2 智算中心网络的行业解决方案

智算中心网络的行业解决方案有很多（如图 3 所示），但从技术路线来看，中兴通讯认为目前主要有增强以太网、改良型 IB 和新型网络 3 种行业解决方案。3 种技术路线各具特点。

1) 增强型以太网

增强型以太网的基本思路是在现有以太网技术的基础上



▲图 3 智算中心网络的行业解决方案演进

针对智算场景的需求开展体系化的优化，以满足智算中心网络的业务需求。目前业界普遍认可的是采用端网协同优化以太网在拥塞控制、时延、抖动以及吞吐方面的性能。该技术路线是阿里、腾讯等头部互联网公司的一致选择，也是中兴通讯智算中心网络的主要产品路线。

2) 改良型IB网络

长期以来，英伟达将IB作为其智算中心网络的主要解决方案，这给业界传递了一种不太准确的信息：IB是最适合智算中心网络的方案。事实上，在实际的智算场景测试中，以太网在带宽、组网规模等方面均不逊于IB^[32]。从技术角度分析，IB网络并非是为当前新兴的AI业务量身定制的，而是传统高性能计算（HPC）市场的方案传承。英伟达宣称的大部分IB技术优势如自适应路由、SHARP在网计算等均是基于无限带宽网络贸易协会（IBTA）标准的私有化改良。随着智算中心网络的发展，IB网络也将不断演进和改良，并在相当长一段时间内维持其在智算中心网络领域的市场份额。

3) 新型网络

由于智算场景对规模和性能的极致追求，行业内开展了很多新型网络技术的探索 and 实验。这类方案的特点是试图最大限度重用现有产业链：在网络拓扑、互联技术方面创新，追求极致性能。但这样会牺牲一定的异厂商互通性。典型的技术方案有开放计算项目（OCP）提出的分布式调度网络（DSF）方案，以及中国移动提出的GSE方案。

4 中兴通讯在智算中心网络中的技术和产品创新

基于对智算中心网络需求和行业趋势的洞察，中兴通讯在坚持核心技术自研的前提下，沿着增强以太网和新型网络两条路线开展了一系列技术和产品的创新，形成了完整的智算中心网络解决方案。

1) 基于自研芯片的智算交换机产品

以全自研芯片为基础，中兴通讯已经形成了完整的智算系列交换机产品（如图4所示），包括盒式59和框式99两大系列。59盒式系列交换机设备单机容量达到12.8T，并将快速迭代到单机51.2T，性能和可编程能力达到业界先进水平。99框式系列交换机设备采用中国性能最高、国际领先的自研分布式芯片，可以提供多达576个400G端口密度，单层网络即可支持万卡以上规模集群的组网需求。

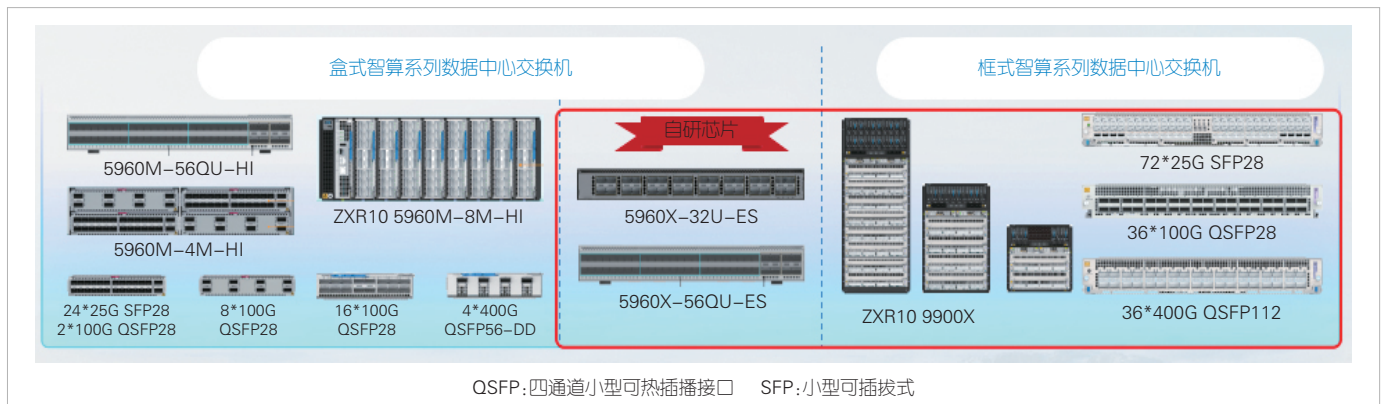
2) 智算中心网络技术和方案创新

在智算中心网络的技术路线上，中兴通讯沿着增强以太网和新型网络方面开展技术和方案创新，形成了差异化的解决方案。

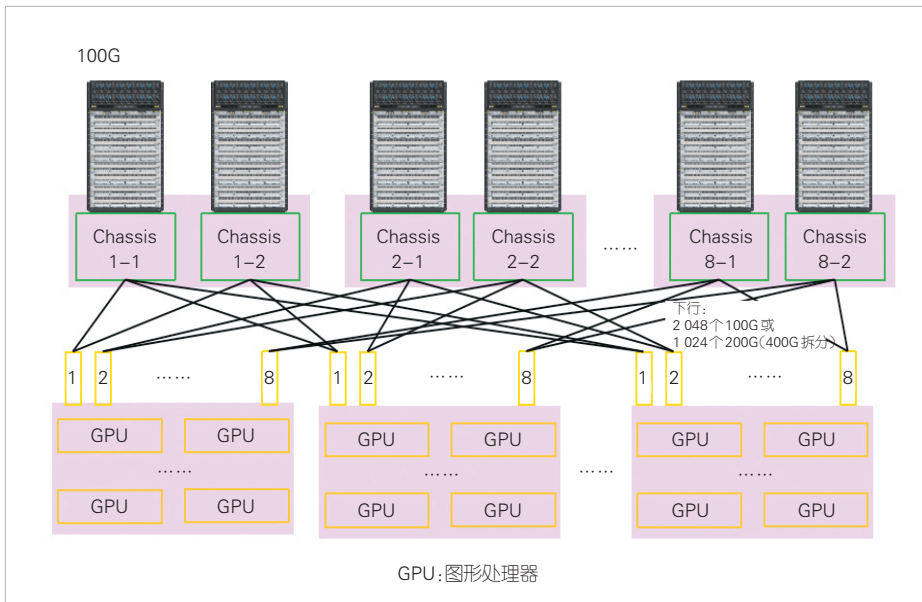
(1) 增强以太网的技术创新实践

通过自研DPU产品和交换机之间的端网协同设计，中兴通讯解决了网络拥塞状态的精细化感知、网络多路径、全局负载均衡、端网统一管控等一系列工程化难题。

中兴通讯在智算网络技术上不断创新，通过轨道负载分担（ZRLB）技术，基于增加交换机入端口（Ingress Port）作为Hash Key的算法，根据业务需求把连接服务器的端口进行Group分组，并基于Group内物理口进行Hash，实现出口流量均衡，提升网络负载均衡效率；通过智能全局负载分担（iGLB）技术，实现“网络控制器+AI调度平台”协同工作，精确掌握每条业务流的带宽诉求，集中计算出每条业务流的最优转发路径，达到整网负载均衡效率最佳；通过端网协同拥塞控制（ENCC）技术，带内遥测扩展链路状态信息，端网协同实现精准拥塞控制、链路故障实时反馈、网络快速发送拥塞通知报文，从而改进流量调度；同时控制器通告整网多路径信息，端侧DPU根据流负载均衡选路，网络按规划路径转发，保证智算业务流量能够充分利用网络多路径资源，提高传输吞吐率，实现端网协同的新型精准拥塞流控。



▲图4 中兴通讯智算系列交换机产品



▲图5 中兴通讯单层多轨双平面组网方案

通过自适应路由通告（ARN）技术，依靠芯片硬件级支持能力，对链路的拥塞、中断故障快速“检测、传递、切换”，实现端到端的路径切换时间小于1ms。

(2) 新型网络的方案创新实践

以自研分布式芯片和对应的业界最高密度框式交换机为基础，中兴通讯采用单层多轨双平面的创新组网方案（如图5所示），提升网络可靠性和网络规模，确保将网络侧的故障影响降到最低，完美支持16K A800集群的组网需求。相比传统方案，该方案网络层次极简，带宽利用率接近100%，大幅度降低光模块的互联成本。

5 结束语

智算中心网络作为支撑AI业务的重要基础设施，需要应用子层、网卡子层、网络子层以及管控子层构成的完整技术栈的方方面面一起协作、创新，以提供超大网络规模、无损、低时延、高吞吐、高可靠以及高可维能力等高性能网络的技术特性。智算中心网络在未来几年将迎来市场和技术的跨越式发展。综合考虑当前AI大模型的发展趋势和中国算力基础设施的现状，支撑百万级集群规模应为智算中心网络的基本要求。但是面向中国单点算力建设规模受限、算力碎片化严重的现状，大数据入算的新场景、跨域分布式训练、高通量数据传输等新需求陆续出现，对网络提出了更高的要求，规模上的量变将带来技术上的质变。网络在支持大模型训练的同时，还需要具备训推一体、支持多租户隔离的网络架构，从而推进大模型更广泛的应用。未来几年，面向AI

场景优化将成为智算中心网络发展的主旋律，行业必须在基础芯片、设备形态、网络架构、网络协议以及应用生态等方面做出更多努力，以进一步推进算侧、端侧和网络侧关键技术的融合发展。中兴通讯相信开放繁荣的智算中心网络生态才是行业的未来，并将持续在该领域做出更多的原创性成果。

参考文献

[1] KODALI R K, PRASAD UPRETI Y, BOPPANA L. Large language models in AWS [C]//Proceedings of 1st International Conference on Robotics, Engineering, Science, and Technology (RESTCON). IEEE, 2024: 112–117. DOI: 10.1109/restcon60981.2024.10463557

[2] YELURI S. Large language models: the hardware connection [EB/OL]. [2024-10-10]. <https://blog.apnic.net/2023/08/10/large-language-models-the-hardware-connection>

[3] TANG Z H, SHI S H, WANG W, et al. Communication-efficient distributed deep learning: a comprehensive survey [EB/OL]. (2020-03-10)[2024-10-06]. <https://arxiv.org/abs/2003.06307>

[4] NVIDIA. NVIDIA spectrum-X network platform architecture [EB/OL]. [2024-10-06]. <https://resources.nvidia.com/en-us-accelerated-networking-resource-library/nvidia-spectrum-x>

[5] CISCO. Evolve your AI/ML network with Cisco silicon one [EB/OL]. [2024-10-06]. <https://www.cisco.com/c/en/us/solutions/collateral/silicon-one/evolve-ai-ml-network-silicon-one.html>

[6] ZHANG Z, LUO L, NING Q, et al. SRNIC: a scalable architecture for RDMA NICs [EB/OL]. [2024-10-03]. <https://www.usenix.org/conference/nsdi23/presentation/wang-zilong>

[7] ZHANG Z L, CAI D Q, ZHANG Y R, et al. FedRDMA: communication-efficient cross-silo federated LLM via chunked RDMA transmission [EB/OL]. (2024-03-01)[2024-10-08]. <https://arxiv.org/abs/2403.00881>

[8] TANG J, WANG X L, DAI H C. Scalable RDMA transport with efficient connection sharing [EB/OL]. [2024-10-05]. <https://ieeexplore.ieee.org/document/10228968>

[9] ZHU Y, ERAN H, FIRESTONE D, et al. Congestion control for large-scale RDMA deployments [EB/OL]. [2024-10-05]. <https://dl.acm.org/doi/10.1145/2785956.2787484>

[10] LI Y, MIAO R, L H H, et al. HPCC: high precision congestion control [EB/OL]. [2024-10-07]. <https://dl.acm.org/doi/10.1145/3341302.3342085>

[11] KUMAR G, DUKKIPATI N, JANG K, et al. Swift: delay is simple and effective for congestion control in the datacenter [EB/OL]. [2024-10-12]. <https://dl.acm.org/doi/pdf/10.1145/3387514.3406591>

[12] WANG W, MOSHREF M, LI Y, et al. Poseidon: efficient, robust, and practical datacenter CC via deployable INT [EB/OL]. [2024-10-11]. <https://www.usenix.org/conference/nsdi23/presentation/wang-weitao>

[13] JOSHI R, SONG C H, KHOOI X Z, et al. Masking corruption packet losses in datacenter networks with link-local retransmission [C]//Proceedings of the ACM SIGCOMM 2023 Conference. ACM, 2023: 288–304. DOI: 10.1145/

- 3603269.3604853
- [14] YAN B, ZHAO Y, XU S, et al. LHCC: low-latency and hi-precision congestion control in RDMA datacenter networks [EB/OL]. [2024-10-06]. <https://ieeexplore.ieee.org/document/10682889>
- [15] LIM H, KIM J, CHO I, et al. FlexPass: a case for flexible credit-based transport for datacenter networks [EB/OL]. (2023-05-08) [2024-10-07]. <https://dl.acm.org/doi/10.1145/3552326.3587453>
- [16] SONG C H, KHOOI X Z, JOSHI R, et al. Network load balancing with In-network reordering support for RDMA [C]//Proceedings of the ACM SIGCOMM 2023 Conference. ACM, 2023: 816-831. DOI: 10.1145/3603269.3604849
- [17] HUANG P, ZHANG X, CHEN Z, et al. LEFT: lightwEight and fast packet reordering for RDMA [EB/OL]. (2024-08-03) [2024-10-08]. <https://dl.acm.org/doi/abs/10.1145/3663408.3663418>
- [18] CHEN C, YE J, GAO Y, et al. HF2T: host-based flowlet fine-tuning for RDMA load balancing [EB/OL]. [2024-10-10]. <https://dl.acm.org/doi/10.1145/3663408.3663410>
- [19] DENG H T, YANG Y, ZHANG M, et al. CAVER: enhancing RDMA load balancing by hunting less-congested paths [EB/OL]. [2024-10-10]. <https://dl.acm.org/doi/10.1145/3672202.3673729>
- [20] DONG J B, CAO Z, ZHANG T, et al. EFLOPS: algorithm and system co-design for a high performance distributed training platform [C]//Proceedings of IEEE International Symposium on High Performance Computer Architecture (HPCA). IEEE, 2020: 610-622. DOI: 10.1109/hpca47549.2020.00056
- [21] KIM J, DALLY W J, SCOTT S, et al. Technology-driven, highly-scalable dragonfly topology [EB/OL]. [2024-10-12]. <https://ieeexplore.ieee.org/document/4556717>
- [22] JOUPPI N P, KURIAN G, LI S, et al. TPU v4: an optically reconfigurable supercomputer for machine learning with hardware support for embeddings [EB/OL]. (2023-04-04) [2024-10-05]. <https://arxiv.org/abs/2304.01433>
- [23] IETF. Routing in fat trees (RIFT) working group [EB/OL]. [2024-10-05]. <https://datatracker.ietf.org/doc/draft-ietf-rift-rift/>
- [24] NVIDIA. Adaptive routing [EB/OL]. [2024-10-05]. <https://docs.nvidia.com/networking-ethernet-software/cumulus-netq-48/Monitor-Operations/Monitor-Adaptive-Routing/>
- [25] OPEN COMPUTE PROJECT 2020. Distributed disaggregated chassis routing system [EB/OL]. [2024-04-10]. <https://www.opencompute.org/documents/ufispace-dcc-routing-system-intro-for-ocp-summit-2020-1-pdf>
- [26] 段晓东, 程伟强, 王瑞雪, 等. 面向新型智能计算中心的全调度以太网技术 [J]. 中兴通讯技术, 2023, 29(4): 57-63. DOI: 10.12142/ZTETJ.202304011
- [27] NVIDIA. SHARP: in-network scalable streaming hierarchical aggregation and reduction protocol [EB/OL]. [2024-10-12]. <https://mug.mvapich.cse.ohio-state.edu/static/media/mug/presentations/20/bureddy-mug-20.pdf>
- [28] NVIDIA. NVIDIA GB 200 NVL72 [EB/OL]. [2024-10-13]. <https://www.nvidia.com/en-us/data-center/gb200-nvl72/>
- [29] POUTIEVSKI L, MASHAYEKHI O, ONG J, et al. Jupiter evolving: transforming google's datacenter network via optical circuit switches and software-defined networking [EB/OL]. [2024-10-12]. <https://dl.acm.org/doi/pdf/10.1145/3544216.3544265>
- [30] WANG W Y, GHOBADI M, SHAKERI K, et al. Rail-only: a low-cost high-performance network for training LLMs with trillion parameters [EB/OL]. [2024-10-13]. <https://ieeexplore.ieee.org/document/10664412>
- [31] RABINOVITSJ D. Opening AI infrastructure [EB/OL]. [2024-10-13]. <https://drive.google.com/file/d/1ud1JZqco2868AvmkNkrA-Axp-74PvwWx/view>
- [32] BROADCOM. Why ethernet reigns supreme over InfiniBand for large-scale networks [EB/OL]. [2024-12-04]. <https://docs.broadcom.com/doc/Unleashing-the-Power-of-AI-ML>

作者简介



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



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



周昆, 中兴通讯股份有限公司研发规划有线总工, 负责AI智算、金融DC等多个业务领域的DCN方案和规划。