Scale-Up网络内GPU卡间互联及OISA技术体系

Intra-Node GPU Interconnection in Scale-Up Networks and OISA Technologies

宋晓勇/Song Xiaoyong，李锴/Li Kai，陈佳媛/Chen Jiayuan
（中国移动通信有限公司研究院，中国 北京 100053）
(Research Institution of China Mobile, Beijing 100053, China)

DOI：10.12142/ZTETJ.202603004
网络出版地址：https://link.cnki.net/urlid/34.1228.TN.20260626.1349.010
网络出版日期：2026-06-26
收稿日期：2026-04-05

摘要：针对图形处理器（GPU）Scale-Up卡间互联，系统梳理了其发展背景、核心需求与现实挑战，总结了从通用总线、私有协议到开放标准的3阶段演进历程及技术趋势。以中国移动提出的全向智感互联架构（OISA）为分析对象，详细阐述了其架构设计、核心技术特征与创新优势。在此基础上，结合大模型推理阶段键值缓存（KV Cache）的存储通信需求，以及超节点场景下Scale-Up互联面临的新挑战，对其未来发展方向进行了展望与探讨，为智算网络体系构建及超节点演进提供了参考。研究结果表明，各协议技术正逐渐趋同，开放互联已成为重要发展态势。OISA依托分层协议栈设计与关键技术创新，构建了高效、可靠、灵活、开放的GPU高速互联体系，为智算服务器向超节点迭代升级提供了关键支撑。

关键词：垂直扩展；Scale-Up；卡间互联；全向智感互联；超节点；智算网络

Abstract: A systematic review is conducted on the background, core requirements, and practical challenges of graphics processing unit (GPU) Scale-Up interconnection, and its three-stage evolution from general-purpose buses and proprietary protocols to open standards, along with corresponding technical trends, is summarized. With the omnidirectional intelligent sensing express architecture (OISA) proposed by China Mobile taken as the analytical object, its architecture design, key technical features, and innovative advantages are elaborated in detail. On this basis, combined with the storage and communication demands of Key-Value (KV) Cache in large-model inference and the emerging challenges faced by Scale-Up interconnection in SuperPod scenarios, its future development directions are discussed and prospected, providing references for intelligent computing network construction and SuperPod evolution. The results indicate that interconnection protocols and technologies are gradually converging, and open interconnection has become a key development trend. Based on hierarchical protocol stack design and key technological innovations, OISA establishes an efficient, reliable, flexible, and open high-speed GPU interconnection system, providing critical support for the iterative upgrade of intelligent computing servers toward SuperPods.

Keywords: vertical expansion; Scale-Up; intra-node GPU interconnection; omni-directional intelligent sensing express architecture (OISA); SuperPod; intelligent computing network

引用格式：宋晓勇, 李锴, 陈佳媛. Scale-Up网络内GPU卡间互联及OISA技术体系 [J]. 中兴通讯技术, 2026, 32(3): 15-21. DOI: 10.12142/ZTETJ.202603004

Citation: Song X Y, Li K, Chen J Y. Intra-node GPU interconnection in scale-up networks and OISA technologies [J]. ZTE technology journal, 2026, 32(3): 15-21. DOI: 10.12142/ZTETJ.202603004

1 Scale-Up卡间互联

1.1 背景

人工智能技术的爆发式发展，推动大语言模型（LLM）和混合专家模型（MoE）向万亿乃至十万亿参数规模演进。在此过程中，模型训练与在线推理对参数、激活值和梯度的交互提出了极高密度、低延迟和强实时性的要求，进而对算力、存储及通信效率构成严峻挑战。单张图形处理器（GPU）的算力和显存容量已无法独立支撑此类大模型的运行，必须依托数据并行（DP）、张量并行（TP）、MoE并行、流水线并行（PP）等分布式并行策略，由多张GPU协同完成计算任务。然而，由于GPU之间的数据搬运速度明显滞后于其计算速度，大量算力因等待数据而空转。因此，内存共享效率与通信效率成为制约多GPU协同性能的关键因素。

智算网络是实现多GPU协同的关键基础设施，其性能直接决定模型训练的迭代周期与推理响应速度。智算网络主

要分为3类，即Scale-Up（垂直扩展）网络、Scale-Out（水平扩展）网络和Scale-Across（跨域扩展）网络[1]。其中，Scale-Up网络专注于单服务器节点或超节点内的GPU高速卡间互联，尤其适配TP、MoE并行等紧耦合任务；这与聚焦同园区内服务器间互联、适配数据并行等松耦合任务的Scale-Out网络不同，也不同于面向跨园区多智算中心广域协同的Scale-Across网络。高性能卡间互联协议及相关技术正是构建Scale-Up网络的核心[2-3]。

1.2 需求与挑战

Scale-Up网络通过超高带宽、极低时延的互联，突破单节点内多GPU的数据交互壁垒与存储访问边界，构建逻辑一体化的“超级GPU”架构，即超节点。该网络性能持续跃升，推动服务器硬件架构不断演进，算力部署向高密度集约化方向升级；超节点逐步成为新一代智算基础设施的主流形态，同时也对卡间互联技术提出了极为严苛的要求[4]。

在带宽方面，卡间互联技术需匹配高带宽内存（HBM）的访问速率，GPU单端口带宽需达400～800 Gbit/s，并在高端场景中向更高速率演进，以满足TP海量参数的即时同步需求。在时延方面，GPU间往返时延（RTT）需维持在微秒级，交换拓扑下的单跳转发时延需控制在百纳秒级，以保障协同操作的实时性。在扩展性方面，该技术需突破传统直连或Cube-Mesh等拓扑的限制，基于交换拓扑实现数百张GPU卡的无阻塞全互联。功能方面，该技术需具备无损传输机制及集合通信加速（CCA）能力，保障整体传输效率，并通过统一内存寻址等消除数据拷贝开销，简化多GPU编程模型。

面向上述核心需求，Scale-Up互联面临多重瓶颈。在技术层面，卡间互联需同步满足带宽、时延与互联规模的极致约束；传统互联协议的演进迭代滞后于GPU及HBM硬件的升级，普遍存在固有性能短板，且改造适配成本高。随着组网规模扩大，传统方案易出现时延累积、流量拥塞及拓扑适配等难题，叠加大模型训练与推理差异化的通信特征，进一步增加了架构兼容与设计难度。当前中国GPU厂商多采用私有互联协议，缺乏专用交换芯片的深度适配，阻碍其向超节点架构演进。若针对单一GPU生态定制研发交换芯片，高额研发成本难以通过市场规模实现有效分摊。同时，中国高速率串行器/解串器（SerDes）设计能力与先进芯片制程工艺供给相对不足，叠加高密度集成场景下的散热与能耗管控压力，抬高了超节点一体化设计门槛与工程实现难度。

综上所述，现有Scale-Up互联技术在性能、生态与工程化层面均存在明显局限，亟需一套开放、高性能的互联协议体系，以支撑下一代智算超节点的发展。

2 卡间互联技术演进

2.1 演进历程

GPU Scale-Up卡间互联的演进可划分为以下3个阶段[5]。

第1阶段：传统通用互联阶段。在此阶段，产业界普遍采用外围组件互连高速（PCIe）总线实现GPU卡间互联。尽管PCIe 6.0及之后版本通过调制编码优化提升了带宽，但其原生设计面向通用外设互联，且采用根复合体（Root-Complex）主从架构，导致其在智算服务器GPU协同场景下存在固有缺陷。具体而言，其带宽与时延难以满足紧耦合并行计算的需求，树状拓扑结构也限制了大规模组网能力，最终成为制约算力释放的瓶颈[6]。

第2阶段：私有协议主导的定制化互联阶段。为突破PCIe架构的性能限制，GPU厂商围绕自身芯片特性推出了专属互联协议。在国际头部厂商中，NVIDIA于2014年发布NVLink[7]，AMD于2017年推出Infinity Fabric Link[8]，Intel于2020年提出X^e Link；中国厂商也同步推进，壁仞科技在BR100产品中采用BLink，摩尔线程在MTT S4000产品中采用MTLink，均实现了GPU全互联。这些私有协议大幅提升了卡间通信性能，单卡带宽可达200～400 GB/s，点对点时延显著降低。然而，除NVIDIA的NVLink配有NVSwitch专用交换芯片外，绝大多数私有协议缺乏交换芯片支持，仅能实现8卡互联，无法支撑Scale-Up网络的规模化扩展需求。

第3阶段：规模化与标准化互联阶段。面向超节点百卡级乃至更高密度的Scale-Up互联需求，产业界加速构建开放标准协议体系。2024年，中国移动率先提出全向智感互联（OISA）技术体系，并于2025年8月发布2.0版本协议规范；同年，AMD主导联合阿里巴巴、Astera Labs、Intel、Meta、Synopsys等企业成立统一加速器互联（UALink）联盟，并于2025年4月发布UALink 1.0协议[9]；Broadcom于2025年4月发布可扩展统一以太网（SUE）初版规范；华为于2025年9月发布灵衢（UB）基础规范[10]；开放计算项目（OCP）联合AMD、Broadcom、Cisco、NVIDIA等厂商于2025年10月成立增强型扩展网络（ESUN）工作组，并于2026年2月发布首版基础规范[11]。

这些新型开放标准互联协议的集中涌现，从根本上打破了单一厂商的生态垄断，实现了GPU与交换芯片之间的互联互通，推动了卡间互联从直连拓扑向基于交换架构的全互联模式转变，显著提升了大规格组网能力与拓扑扩展性。

2.2 演进趋势

随着ScaleUp互联进入开放化、标准化阶段，各类新型

卡间互联协议虽然在报文格式与技术细节上存在差异，但其整体演进呈现出清晰的共性趋势。

1）协议轻量化和高效化。面向GPU高频细粒度通信与内存语义访问模式，各协议以实现超高带宽、超低时延为核心目标，持续精简协议开销，最大化提升有效带宽利用率，支持低时延的Load/Store访问，从而充分释放GPU算力潜能。

2）互联拓扑向基于交换的无阻塞全互联升级。如图1所示，GPU卡间互联拓扑从传统PCIe树状架构、私有协议的点对点直连，转向以专用ScaleUp交换芯片为核心的全互联架构，可支持百卡至千卡级高密度组网，单跳转发时延控制在百纳秒级，从而满足超节点性能及大规模扩展需求。

3）传输机制趋于无损化。各协议普遍采用基于信用的流量控制（CBFC）和基于优先级的流量控制（PFC），并结合链路级重传机制，实现无丢包、高可靠传输，以满足大模型训练对通信稳定性的严苛要求。

4）物理层向高速化方向演进。单端口速率由200 Gbit/s、400 Gbit/s向800 Gbit/s、1.6 Tbit/s升级，高速串行器/解串器（SerDes）与轻量级前向纠错（FEC）成为标配。

5）生态开放化与多厂商兼容。开放标准协议已成为主流，支持跨厂商GPU与交换芯片互联互通，降低供应链依赖与部署成本，推动ScaleUp技术走向标准化与产业化。

3 全向智感互联OISA技术体系

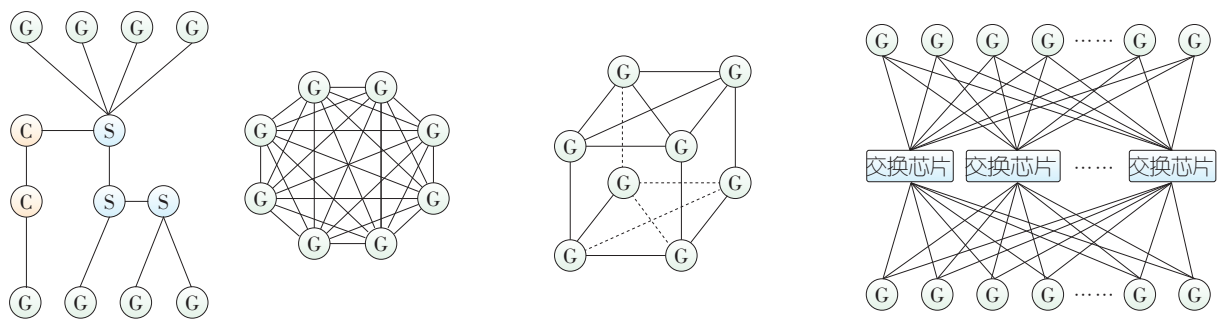
3.1 总体架构

OISA由中国移动于2024年提出，同年相继发布1.0和1.1版本协议规范，并于2025年8月发布2.0版本协议规范[12]。OISA旨在构建高效、智能、灵活、开放的卡间互联体系，基于交换拓扑建立GPU间的对等通信架构，通过原生支持GPU内存语义访问，实现高带宽、低时延的大规模ScaleUp通信，为数据密集型AI应用（如大模型训练、推理及高性能计算）提供支撑。

如图2所示，OISA协议栈采用事务层、数据层和物理层三级分层架构，各层级功能解耦、协同配合，共同支撑高速卡间互联传输。

事务层作为协议栈顶层，直接对接GPU片上网络，主要完成上层访问事务的协议封装与解析，同时集成多模内存访问、集合通信加速、报文聚合分解、即时反馈等核心能力。该层兼容两种内存语义访问模式：一是基于Load/Store指令的同步内存语义（指令直驱模式），二是基于直接内存访问（DMA）的异步内存语义。同时，该层明确了内存栅栏操作要求，以保障多GPU内存访问的一致性与正确性，适配不同业务粒度与性能诉求。数据层位于中间，依托智能状态感知、多种流量控制和链路级重传等关键技术，为上层事务交互提供高效、可靠的传输支撑。物理层采用以太兼容架构，分为逻辑子层（PCS/PMA）与电气子层（PMD），在兼容IEEE 802.3标准、保留生态通用性的基础上，优化高速SerDes、轻量化编解码等底层技术，针对超节点高密度组网场景完成定制化性能优化，实现高带宽、低损耗的物理链路传输。

为适配多样化互联场景，OISA定义了3种工作模式：极致性能访问模式、智能感知访问模式与集合通信加速



（a）PCIe互联拓扑　（b）8卡全互联拓扑　（c）8卡Cube-Mesh互联拓扑　（d）基于交换芯片的全互联拓扑

CPU：中央处理器　GPU：图形处理器　PCIe：外围组件互连高速

图1　GPU卡间互联拓扑

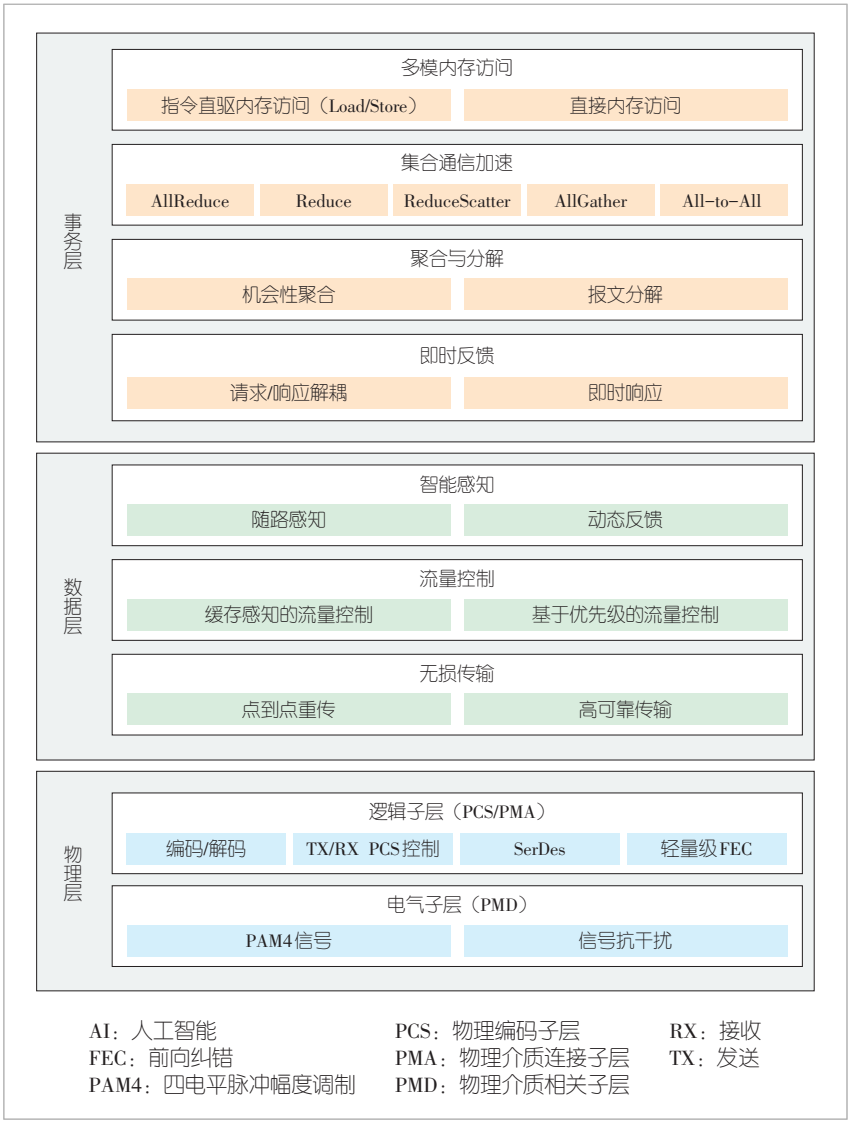


图2　OISA 2.0协议栈

（CCA）模式。其中，极致性能模式面向低时延内存语义访问，通过简化转发流程实现线速处理；智能感知模式支持路径状态采集与动态反馈，实现端到端协同传输优化；集合通信加速模式支持将AllReduce等集合通信操作卸载至交换芯片完成，从而提升分布式训练效率。

3.2 关键技术

OISA构建了一套适配高性能计算与AI场景的高效互联架构。该架构不仅依赖流量控制和重传等传输优化机制，更强调GPU计算单元与卡间互联网络之间的协同，并针对计算侧的业务特性与互联传输痛点进行专门设计。

3.2.1 事务报文动态机会聚合机制

针对推理场景中高频细粒度的通信需求（典型事务大小为64 B~256 B），OISA在事务层设计了报文动态聚合机制，以降低小型事务传输中协议头部开销的占比。

从技术原理来看，在传统的单事务独立封装模式下，64 B载荷所对应的事务层数据包（TLP）头部开销占比高达20%~40%；对于无载荷的请求或响应类报文，该占比进一步上升。OISA采用动态机会聚合策略，在满足时延约束的前提下，将同目的地址、同类型的多个小型事务报文聚合为单个TLP进行传输，从而有效提升传输效率。

如图3所示，聚合报文将固定协议开销分摊至多个子事务，从而显著提升载荷率与带宽利用率。同时，接收端配备硬件级并行解析单元，依托预定义的载荷分隔标识和流水线处理架构，实现报文的快速解包，确保整体传输效率。

3.2.2 集合通信加速

OISA以“计算卸载”重构分布式通信架构，将AllReduce、ReduceScatter、All-Gather、All-to-All等大模型训推场景高频调用的集合通信计算任务，从GPU侧卸载至交换芯片集成的专用计算单元执行。

如图4所示，以AllReduce操作为例，传统直连拓扑通常采用参数GPU集中规约或环形AllReduce等方案。在此模式下，GPU除承担梯度计算任务外，还需参与中间结果的收发、聚合与同步过程。这不仅消耗片上算力资源，亦因多轮端到端数据交互而产生高昂的通信开销。即便借助交换拓扑扩展了互联规模，缺乏计算能力的普通交换芯片仍仅承担报文转发功能，并未改变GPU主导集合通信计算的本质，中间结果反复传输所带来的带宽占用与延迟瓶颈仍未得到有效缓解。

OISA则将AllReduce等集合通信的计算任务直接卸载至交换芯片内部完成。各GPU仅需将本地梯度发送至交换芯片，交换芯片的专用计算单元在报文转发的同时，同步完成求和、求均值等聚合运算，并将最终结果返回至通信组内各节点。该模式从根本上削减了GPU间的中间数据交互量，大幅缩短了数据传输路径与集合通信的整体完成时间。与此同时，GPU算力得以从通信任务中释放，专注于模型训练，从而显著提升分布式集群的整体计算效率。

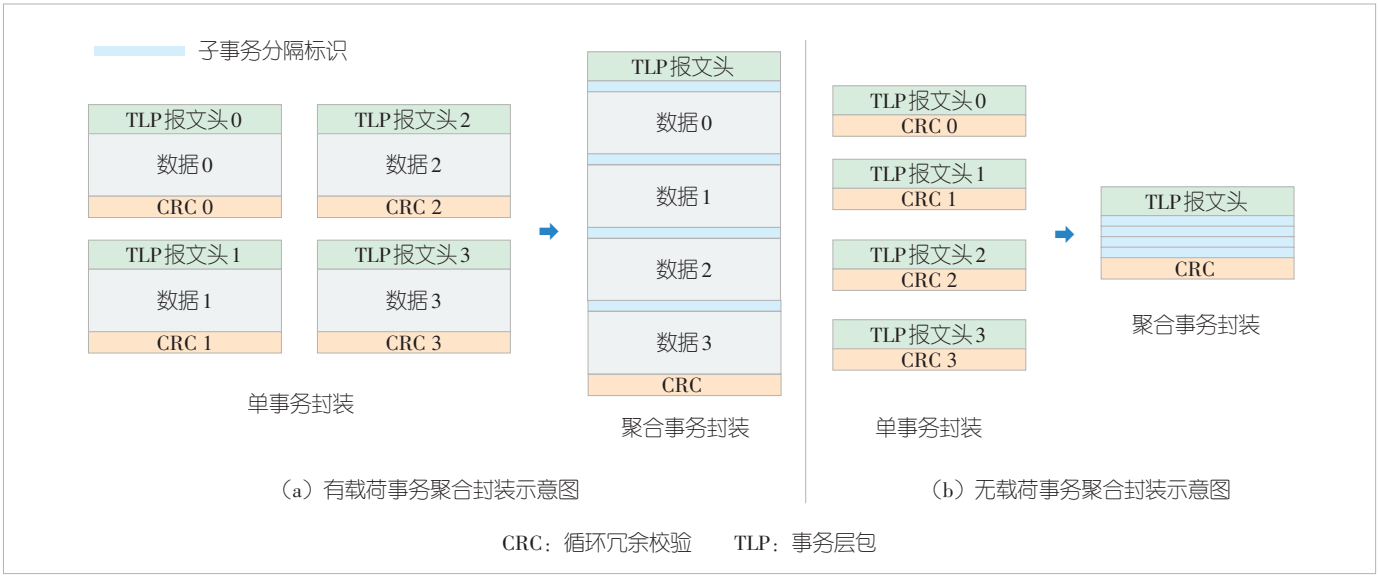


图3 聚合报文格式示意图

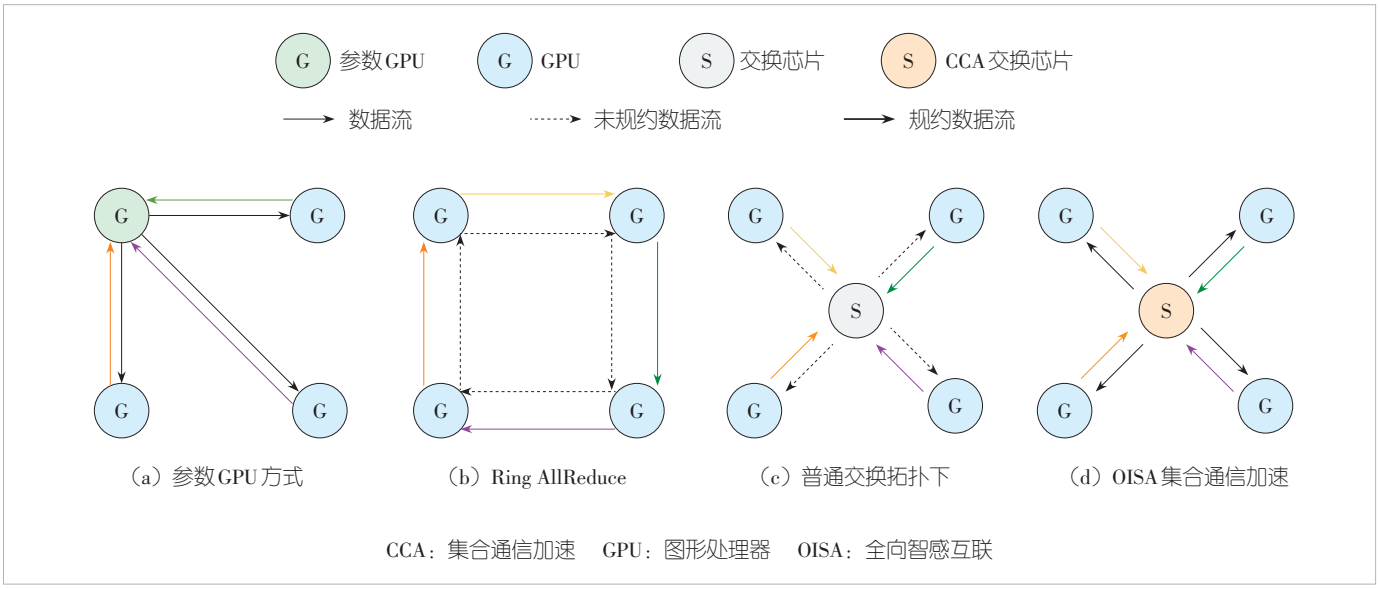


图4 集合通信加速示意图（以规约操作为例）

3.2.3 稀疏感知加速机制

针对MoE模型训练推理中因跨GPU通信突发性强、分布不均匀、小包占比高等特征所引发的链路拥塞与时延劣化问题，OISA引入了稀疏感知加速机制。该机制通过在事务报文头部嵌入稀疏感知标签，使GPU能够标记当前激活的专家状态；交换芯片则据此仅对活跃数据进行聚合，并动态优化转发调度策略。同时，结合报文聚合技术降低协议开销，有效缓解非对称流量模式所导致的性能衰减，从而充分适配超大参数量MoE模型对高效通信的严苛需求。

3.2.4 智能感知

OISA的智能感知模式通过“发送端主动探测”与“接收端主动告警”两种机制实现随路感知和动态反馈，为传输控制提供精准依据，有效规避拥塞瓶颈。

如图5所示，发送端主动探测机制由发送端GPU在报文中嵌入轻量化感知标签，沿数据传输路径随路采集各节点状态信息，并记录潜在瓶颈点。当接收端检测到标签所携带的状态信息达到预设阈值（例如队列占用率超过70%）时，即向发送端反馈响应信息。发送端收到反馈后，识别并判断拥塞点的具体位置及拥塞程度，进而通过降低发送速率或调整

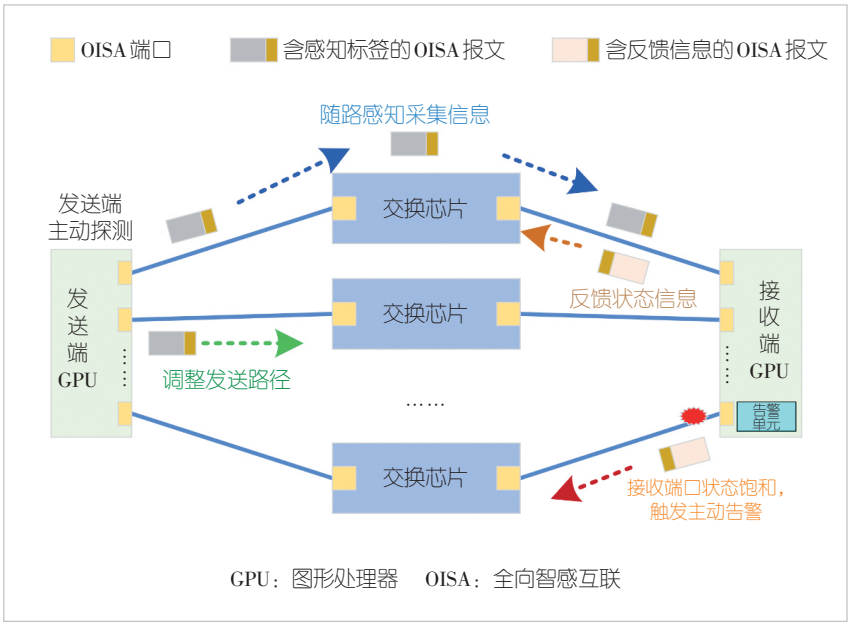


图5　智能感知示意图

转发路径等手段，防止拥塞状况进一步恶化。

接收端主动告警机制则从接收侧提供兜底保障。当接收端GPU的接收性能出现持续下降时，该机制主动向发送端反馈状态信息，以应对发送端主动探测未能覆盖的接收端侧拥塞风险。

上述两种机制协同互补，使发送端能够全面感知网络链路与接收端的实时状态，并作出快速响应，从而从源头规避因局部链路拥塞所引发的全局传输瓶颈，切实保障大规模Scale-Up集群内数据传输的低时延与高稳定性。

4 未来发展方向

尽管Scale-Up网络在互联架构与协议技术方面已日趋成熟，但伴随大模型长文本推理、多模态融合、超大规模MoE的演进，现有卡间互联体系仍面临一系列新挑战，亟需从内存架构、传输介质、系统协同等层面实现系统性突破。

随着GPU单卡算力和互联规模的持续提升，内存访问效率已逐渐成为制约模型部署效率和系统整体性能提升的核心瓶颈[13]。在当前万亿参数大模型在线推理阶段，键值缓存（KV Cache）占用显存比例可超过90%，且其占用量随对话轮次与上下文长度的增长呈线性递增趋势，导致单卡本地显存难以支撑高并发长序列业务场景。与此同时，传统存储架构资源碎片化、跨卡共享能力不足、显存分配不均等问题，易引发局部资源过载与频繁数据置换，大幅降低推理吞吐并增大会话时延。为此，Scale-Up互联的范畴应突破当前GPU卡间互联的局限，延伸至内存池、专用存储节点等组件。通过构建硬件级的一致性缓存域，GPU可直接经由高速总线访问池化内存中的数据，实现计算资源与存储资源的解耦共享，从而有效满足海量KV Cache分布式调度的刚性需求。

在传输介质层面，超节点高密度部署需求与电互联物理极限的共同作用，正加速推动光互联技术的升级演进[14-15]。传统电互联已逐渐逼近物理极限，在高密度布线环境下存在信号衰减严重、传输距离受限等问题，难以支撑跨机箱、跨机柜的无阻塞互联需求。相比之下，光互联具备高带宽密度、低传输损耗与强抗干扰能力等优势，成为突破电互联瓶颈的关键技术路径。未来，共封装光学（CPO）、光输入输出（OIO）片上光互联等光电融合技术的应用，可进一步提升硬件集成度，降低传输功耗；同时，结合光交换技术支持的动态拓扑重构能力，可灵活适配MoE模型非均匀突发流量特征，构建稳定高效的全光互联底座，从而支撑超节点在远距离、大规模场景下的无阻塞通信需求。

在系统协同层面，超节点的规模化部署对全局资源协同提出了更高要求，Scale-Up互联正逐步成为统筹算力与存储资源高效协同的核心枢纽。传统计算与网络相分离的设计范式，易导致互联带宽与GPU算力、存储带宽之间不匹配，进而制约系统整体效能的释放。因此，后续研究应聚焦全栈协同设计理念，在架构设计阶段统筹多组件间的性能适配，充分发挥交换芯片在集合通信加速与流量调度方面的能力，以提升资源利用效率；同时，应加快推进开放互联协议的标准化进程，着力破除跨厂商硬件适配的壁垒，推动超节点从硬件堆叠式部署向一体化全局协同的智算架构演进。

5 结束语

随着大模型参数量与算力需求的持续攀升，基于Scale-Up互联架构的超节点已成为新一代高密度智算基础设施的核心形态，相关产业共识正逐步形成。当前，全球虽存在多种新型卡间互联协议并行发展的格局，但整体技术路线与演进方向已呈现收敛态势，基于新协议的超节点正加速研发进程，并稳步推进商用部署与规模化落地。

OISA立足国产自主技术路线，构建了高性能、开放可扩展的GPU Scale-Up互联技术体系，有效突破了传统互联在带宽、时延与组网规模等方面的瓶颈，能够精准适配大模型训练推理等数据密集型AI业务的严苛需求。目前，OISA

已完成2.0版本核心协议规范的制定及关键功能验证，硬件研发与产业适配工作正稳步推进。

面向未来发展的持续诉求，OISA将秉持长期演进策略，不断迭代完善。在架构层面，持续拓展互联边界，构建内存共享池，并优化报文结构与流量调度机制，进一步提升集群整体性能与运行稳定性；在传输层面，深度融合CPO/近封装光学（NPO）光互连与Chiplet等先进封装技术，构建基于光互连的OISA超节点，进一步降低通信时延与系统功耗。与此同时，OISA将持续强化多厂商硬件的兼容性与生态开放性，积极推动协议标准化与产业化落地，为中国超节点的自主可控建设以及大模型技术的创新突破构筑坚实互联底座，助力智算基础设施向更高性能、更优能效与更广兼容的方向持续演进。

参考文献

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

[2] Li A, Song S L, Chen J Y, et al. Tartan: evaluating modern GPU interconnect via a multi-GPU benchmark suite [C]//Proceedings of IEEE International Symposium on Workload Characterization (IISWC). IEEE, 2018: 191-202. DOI: 10.1109/IISWC.2018.8573483

[3] 崔晨, 吴迪, 陶业荣, 等. 多GPU系统的高速互联技术与拓扑发展现状研究 [J]. 航空兵器, 2024, 31(1): 23-31. DOI: 10.12132/ISSN.1673-5048.2023.0138

[4] 华为技术有限公司, 中国信息通信研究院. 超节点发展报告 [R]. 2025

[5] Song X Y, Zhou D Y, Li K, et al. Survey of intra-node GPU interconnection in scale-up network: challenges, status, insights, and future directions [J]. Future Internet, 2025, 17(12): 537. DOI: 10.3390/fi17120537

[6] Das Sharma D. PCI-express: evolution of a ubiquitous load-store interconnect over two decades and the path forward for the next two decades [J]. IEEE circuits and systems magazine, 2024, 24(2): 47-61. DOI: 10.1109/MCAS.2024.3373556

[7] Foley D, Danskin J. Ultra-performance pascal GPU and NVLink interconnect [J]. IEEE micro, 2017, 37(2): 7-17. DOI: 10.1109/MM.2017.37

[8] Schieffer G, Shi R M, Markidis S, et al. Understanding data movement in AMD multi-GPU systems with infinity fabric [C]// Proceedings of SC24-W: Workshops of the International Conference for High Performance Computing, Networking, Storage and Analysis. IEEE, 2024: 567-576. DOI: 10.1109/scw63240.2024.00079

[9] UALink Consortium. UALink consortium releases the ultra accelerator Link 200G 1.0 specification [EB/OL]. (2025-04)[2026-04-09]. https://ualinkconsortium. org/wp-content/uploads/2025/04/UALink-1.0-Specification-PR_FINAL.pdf

[10] 华为. 灵衢基础规范 [R]. 2025

[11] Open Compute Project. The OCP ESUN 1.0 specification has been released [EB/OL]. (2026-03-10) [2026-04-09]. https://www. opencompute. org/blog/the-ocp-esun-10-specification-has-been-released

[12] 中国移动. 全向智感互联OISA 2.0技术规范 [R]. 2025

[13] Liu F X, Zhang Q H, Shen H J, et al. HyperOffload: graph-driven hierarchical memory management for large language models on SuperNode architectures [PP/OL]. arXiv [2026-04-09]. https://arxiv.org/abs/2602.00748

[14] Saber M G, Jiang Z P. Physical layer standardization for AI data centers: challenges, progress, and perspectives [J]. IEEE network, 2026, 40(2): 147-155. DOI: 10.1109/MNET.2025.3557812

[15] Liu Y H, Cai Z Z, Chen Y X, et al. InfiniteHBD: building datacenter-scale high-bandwidth domain for LLM with optical circuit switching transceivers [EB/OL]. (2025-02-07)[2026-04-09]. https://arxiv.org/abs/2502.03885

作 者 简 介

宋晓勇，中国移动通信有限公司研究院技术专家；主要从事Scale-Up网络、GPU卡间互联协议及架构、硬件网络加速等方面的研究工作。

李锴，中国移动通信有限公司研究院技术经理，主任研究员；主要从事通算/智算领域的协议、芯片、部件及整机等方面的研究工作。

陈佳媛，中国移动通信有限公司研究院技术经理，主任研究员；主要从事智算网络、多样性算力、新型智算、GPU卡间互联等方面的研究工作。