

面向新型智能计算中心的全调度以太网技术



Global Scheduling Ethernet for New Intelligent Computing Center

段晓东/DUAN Xiaodong, 程伟强/CHENG Weiqiang,
王瑞雪/WANG Ruixue, 王雯萱/WANG Wenxuan

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

DOI: 10.12142/ZTETJ.202304011

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20230724.1702.010.html>

网络出版日期: 2023-07-25

收稿日期: 2023-06-08

摘要: 智能计算中心网络作为智能计算中心的连接底座, 需要具备高性能、低时延的通信能力。一旦网络性能不佳, 就会严重影响分布式训练的效果。智能计算中心网络体系是一个多要素融合的复杂系统, 依赖于智能计算业务、网络设备、交换芯片、网卡、仪表等上下游产业协同创新。提出一种新型全调度以太网(GSE)技术架构, 在最大限度地兼容以太网生态链的前提下, 基于报文容器(PKTC)转发、负载均衡机制以及基于报文容器的动态全局调度队列(DGSQ)全局调度技术, 构建超大规模、超高带宽、超低时延、超高可靠的智能计算中心网络, 助力AI产业发展。

关键词: 人工智能生成内容; 智算中心网络; 全调度以太网; 报文容器; 动态全局调度队列

Abstract: The intelligent computing center network, as the connection base of the intelligent computing center, needs to have high performance and low latency communication capability. Once the network performance is poor, it can seriously affect the effect of distributed training. An intelligent computing center network system is a multi-element integration of complex systems, relying on intelligent computing services, network equipment, switching chips, network cards, instruments, and other upstream and downstream industry collaborative innovation. A new global scheduling Ethernet (GSE) technology architecture is proposed to build an ultra-large scale, ultra-high bandwidth, ultra-low latency, and ultra-reliable intelligent computing center network with maximum compatibility with the Ethernet ecosystem, innovative forwarding and load balancing mechanism based on packet containers (PKTC) and dynamic global scheduling queue (DGSQ) global scheduling technology based on packet containers to help the development of AI industry.

Keywords: artificial intelligence-generated content; intelligent computing center network; GSE; packet container; dynamic global scheduling queue

引用格式: 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

Citation: DUAN X D, CHENG W Q, WANG R X, et al. Global scheduling Ethernet for new intelligent computing center [J]. ZTE technology journal, 2023, 29(4): 57-63. DOI: 10.12142/ZTETJ.202304011

1 AI业务与智能计算中心产业的发展

1.1 AI业务发展趋势

人工智能(AI)业务发展经历了漫长的历程。20世纪50年代,人们开始尝试模拟人脑的神经网络来解决计算机视觉和语音识别的问题。但由于当时无法解决神经网络计算复杂度高和可解释性差的问题, AI技术进入了“寒冬”。2012—2017年, Hinton等提出卷积神经网络, 大大推动计算机视觉和深度学习的发展。同时, 基于深度学习的AlphaGo战胜围棋世界冠军, 进一步点燃人们在深度学习领域探索的热情与信心。2017—2022年, 基于大型神经网络的Transformer架构出现, 该模型可以更好地捕捉序列之间的依赖关

系, 开启了基于深度学习的AI新时代。2022年11月, OpenAI公司开发的大规模智能语言模型ChatGPT横空出世。ChatGPT结合了GPT-3.5和GPT-4系列的大型语言模型, 展现出惊人的语言能力^[1]。该模型深入各个领域, 在引爆全球科技领域的同时, 推动AI产业全面进入大模型时代。因此, ChatGPT的出现具有跨时代的意义。

近年来, 随着算力经济的高速发展^[2], AI业务在自动驾驶、语音识别和自然语言处理等领域取得了许多重大成就, 并涌现出人工智能即服务(AIaaS)和模型即服务(MaaS)两种新型服务模式。当前, 教育、医疗、智慧城市和智能制造等行业迫切需要AI赋能, 例如: 华为云、百度云、阿里云和腾讯云等提供AIaaS的企业均为用户提供高品质的人工

智能服务。MaaS拥有经过大量数据集训练和优化的模型，可为用户提供图像识别、自然语言处理、预测分析和欺诈检测等服务。

为推动AI业务的发展，中国陆续给予政策方面的扶持和激励，特别是东数西算工程的全面启动，给AI大模型在智能计算（后文简称为“智算”）中心的快速发展注入强大的助推剂^[3]。AI大模型的参数量呈指数级增长，有力地驱动了“大模型”向“超大模型”演进。与此同时，智算规模和智算需求也呈指数级增长。预计截至2030年，智算占比将达到70%，AI技术将广泛落地，中国将迎来智算中心建设的热潮。

1.2 智算中心产业发展趋势

为加速智能经济发展和产业数字化转型，智算中心作为一种新的关键性信息基础设施进入公众视野。智算中心既不同于超算中心，也不同于互联网企业和运营商的云计算中心。智算中心既要借鉴超算中心分布式集群计算架构，以支持超大规模、复杂度高及多样性的数据处理，又要参照云计算服务模式，采用统一的架构和统一的应用程序编程接口（API），以屏蔽底层技术细节，降低使用门槛，向不同行业提供普适且灵活多样的智算服务。

随着业内领军企业竞相推出千亿、万亿级参数量的大模型，以图形处理器（GPU）、神经网络处理器（NPU）为代表的AI算力设施迅猛发展，使得智算中心底层GPU算力部署规模达到万卡级别。基于数据并行、模型并行的分布式训练成为处理超大模型和超大数据集的关键手段。智算中心集群算力与GPU算力、节点数量、线性加速比、有效运行时间等呈正相关，需要计算、存储和网络资源的协同设计，具体表现在以下几个方面：在计算方面，单机算力无法支撑海量训练数据，需要将计算任务切分到单机级别，以并行计算的集群架构方式提供算力服务；在存储方面，为突破计算节点中大量密集数据存取带来的算力瓶颈，搭建机械硬盘（HDD）、固态硬盘（SSD）、存储类内存（SCM）等异构存储集群，以降低数据访问时延；在网络方面，构建连接中央处理器（CPU）、GPU、存储等异构算力资源的总线级、高性能无阻塞交换网络，以提升网络通信性能和稳定性；在机房建设方面，提前规划“风火水电”等基础设施，引入液冷系统，实现低电源使用效率（PUE）数据中心的高能效利用。由此可见，传统智算中心正在向新型智算中心演进。

面向智能计算业务的发展，新型智算中心围绕“算、存、网、管、效”五大核心技术全面升级，以提升GPU集群算力，打造多元融合存储，构建高速无损网络，管控异构

算力池化，以高效节能控制为目标，构建标准统一、技术领先、兼容开放的智算底座。

2 智算中心网络演进趋势与挑战

2.1 智算中心网络关键特征

随着GPU高速发展和算力需求的激增，算力中心正向集约化方向发展，数据中心从“云化时代”转向“算力时代”。在传统云数据中心中，传统的计算处理任务或离线大数据计算任务以服务器或虚拟机（VM）为池化对象，网络负责提供服务器或VM之间的连接，并聚焦业务部署效率及网络自动化能力；而智算中心是服务于人工智能的数据计算中心，以GPU等AI训练芯片为主，并以提升单位时间、单位能耗下的运算能力和质量为核心诉求，为AI计算提供更大的计算规模和更快的计算速度。传统数据中心通过CPU来执行计算任务，且网络带宽需求为10~100 Gbit/s，并通过使用传输控制协议（TCP）来完成内存数据的读取；而智算中心网络主要用于承载AI训练业务，其GPU算力与CPU相比拥有更高的计算性能，且网络带宽需求为100~400 Gbit/s（甚至达到800 Gbit/s），并可以通过远程直接内存访问（RDMA）来减少传输时延。由于RDMA网络对于丢包异常敏感，0.01%的丢包率就会使RDMA吞吐率变为0，因此大模型训练的智算中心网络需要缩短迭代过程中通信传输数据的时间，降低通信开销，从而减少GPU的计算等待，提升计算效率。综上所述，零丢包、大带宽、低时延、高可靠是智算中心网络最为关键的特征。

2.2 智算中心网络面临的挑战

与传统数据中心不同，智算中心主要用于承载AI模型训练业务，其通信流量主要具备周期性、流量大、同步突发等特点。在大模型训练过程中，通信具有非常强的周期性，且每轮迭代的通信模式保持一致。在每一轮的迭代过程中，不同节点间的流量保持同步，同时流量以on-off的模式突发式传输。以上通信流量的特点对智算中心网络提出了3个需求：

1) 高接入带宽是基础。大模型训练对带宽比较敏感。网络对通信影响最大的是序列化时延，网络通信质量主要取决于有效带宽。但由于网络交换的时间占比不高，静态时延对模型训练效率影响不大。

2) 网络级负载均衡是关键。保证通信的有效带宽是模型训练的关键因素之一。负载均衡技术是保证有效带宽的关键。集合操作通信的完成时间由最慢节点的完成时间决定。

在无阻塞网络中，若链路负载不均衡，则会导致冲突流有效带宽下降，冲突流的序列化时间增加。

3) 高健壮网络是保障。网络持续高可用、故障业务无中断是分布式系统运行的基础。在高健壮网络中，链路故障时网络会达到亚毫秒级的自动收敛，降低了网络故障对网络拥塞的影响。

如今，基于融合以太网承载远程直接内存访问（RoCE）协议的智算中心网络，通常采用五元组哈希实现链路负载均衡技术，以及基于优先级的流量控制（PFC）、显式拥塞通告（ECN）协议实现网络无损，该方案对智算中心网络提出4个挑战：

挑战1：传统基于逐流的等价多路径路由（ECMP）负载均衡技术在流量数小的情况下会失效，导致流量在交换网络发生极化，链路负载不均。当智算中心网络中存在大象流时，很容易发生多个流被散列到相同的路径上的情况，从而导致链路过载，造成某个物理链路负载过大，甚至会出现拥塞而导致报文丢弃。

挑战2：随着网络规模的不断提升，报文交换方式由单网络节点内实现到单网络节点间多跳实现转变，各节点间也从松耦合关系变化为联合转发。业界通过Clos架构搭建大规模分布式转发结构来满足日益增长的转发规模需求。在该架构下，各节点分布式运行和自我决策转发路径导致无法完全感知全局信息和实现最优的整网性能。

挑战3：当前流量进入网络时，在不考虑出端口转发能力的情况下，流量会以“推”的方式进入网络。分布式训练的多对一通信模型产生大量In-cast流量，造成设备内部队列缓存的瞬时突发而导致拥塞甚至丢包，造成应用时延的增加和吞吐的下降。PFC和ECN都是拥塞产生后的事后干预的被动拥塞控制机制，它们无法从根本上避免拥塞。

挑战4：AI训练网络是一个封闭的专用网络，针对训练效率，通过Underlay直接承载AI训练任务，不再划分Overlay平面，使传统SDN能力失效。同时，传统的智能流分析技术已无法满足高性能无损网络隐患识别、故障预测和闭环等运维可视

化需求。

2.3 智算中心网络的演进趋势

综合当前所面临的挑战，未来智算中心网络将向3个方向进行演进：一是从“流”分发到“包”分发演进，即通过提供逐报文容器动态负载均衡机制，消除哈希极化问题，实现单流多路径负载分担，提升有效带宽，降低长尾时延；二是从“局部”决策到“全局”调度演进，即实现全局视野的转发调度机制，并实现集中式管理运维、分布式控制转发，提高网络可用性；三是从“推”流到“拉”流演进，即从被动拥塞控制向依赖“授权请求”和“响应机制”的主动流控转变，最大限度地避免网络拥塞产生，同时需要引入全局集中式管理系统，提升网络自动化及可视化能力。

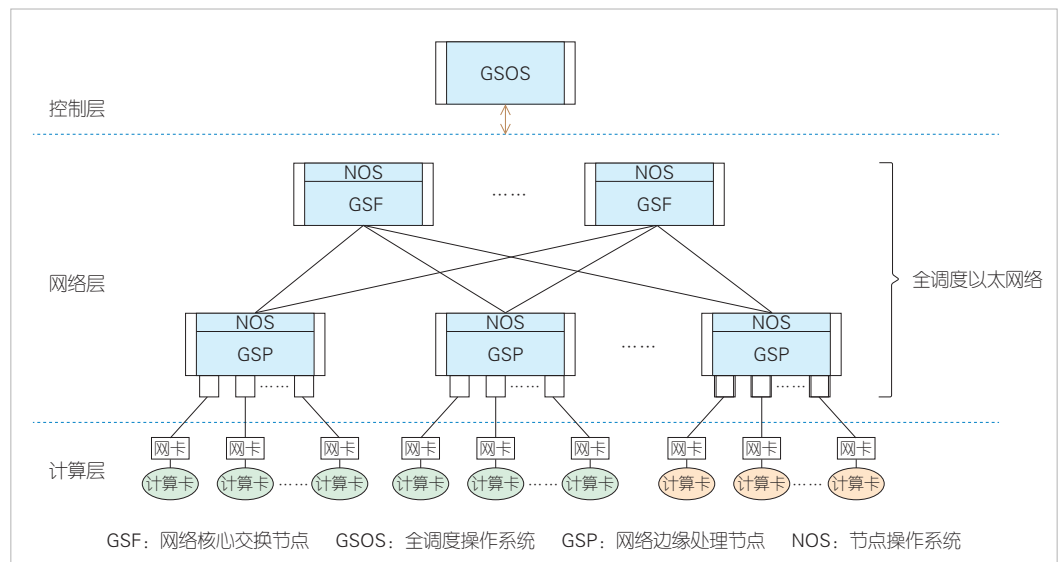
基于以上面向未来智算中心的三大演进方向，我们创新性地提出一种全调度以太网（GSE）技术方案，打造无阻塞、高带宽、低时延、自动化的新型智算中心网络，助力AIGC等高性能业务快速发展^[4]。

3 新型GSE架构体系

3.1 GSE架构介绍

为打造无阻塞、高带宽、低时延的高性能网络，GSE架构应运而生，如图1所示。该架构主要包括计算层、网络层和控制层3个层级，包含计算节点、网络边缘处理节点（GSP）、网络核心交换节点（GSF）及全调度操作系统（GSOS）4类设备^[4]。

1) 控制层：包含全局集中式GSOS，以及GSP和GSF设



▲图1 全调度以太网（GSE）技术体系分层架构

备端分布式节点操作系统 (NOS)。其中, 集中式 GSOS 用于提供网络全局信息, 实现基于全局信息编址、日常运维管理等功能; 设备端 NOS 可实现动态负载均衡、动态全局调度队列 (DGSQ) 调度等分布式网络管控功能。

2) 网络层: GSE 网络主要实现 GSP 和 GSF 协同, 构建出具备全局流量调度、链路负载均衡、流量精细反压等技术融合的交换网络。其中, Fabric 部分可支持二层 GSF 扩展, 以满足更大规模组网需求。

3) 计算层: 即 GSE 网络服务层, 包含高性能计算卡 (GPU 或 CPU) 及网卡, 初期将计算节点作为全调度以太网边界, 仅通过优化交换网络能力提升计算集群训练性能。未来计算将与网络深度融合, 以进一步提升高性能计算能力。

GSE 3 层架构涉及计算节点、GSP、GSF 及 GSOS 4 类设备, 各设备分工如下:

1) 计算节点: 即服务器侧的计算卡、网卡, 提供高性能计算能力。

2) GSP: 即网络边缘处理节点, 用以接入计算流量, 并对流量做全局调度; 流量上行时具备动态负载均衡能力, 流量下行时具备流量排序能力。

3) GSF: 即网络核心交换节点, 作为 GSP 的上一层级设备, 用于灵活扩展网络规模, 具备动态负载均衡能力, 以及反压信息发布能力。

4) GSOS: 即全调度操作系统, 提供整网管控的集中式网络操作系统能力。

换和组包交换。

1) 切包交换: 核心思想是在网络入口将数据包逻辑切分成若干个信元, 将属于同一个数据包的信元调度到不同路径进行传输, 在网络出口侧对信元进行排序及重组, 如图 2 所示。该方式可充分利用多路径交换能力, 最大程度实现链路负载均衡。但在高带宽演进趋势下, 由于被切分后的信元长度短, 信元头部开销带来较多的带宽损耗, 且极高的均衡调度频率对硬件有较高的要求。

2) 逐包交换: 核心思想是不对数据包进行处理, 直接通过轮询或权重等机制将数据包发往不同路径进行传输, 在网络出口侧对报文进行排序, 如图 3 所示。该方式不存在额外的报文开销, 也无需高频的均衡调度周期。但由于数据包长度分布连续, 难以准确根据已发往每条路径的数据包总数据量来实现均衡负载, 链路负载均衡性差, 易受微突发影响导致网络拥塞甚至丢包。

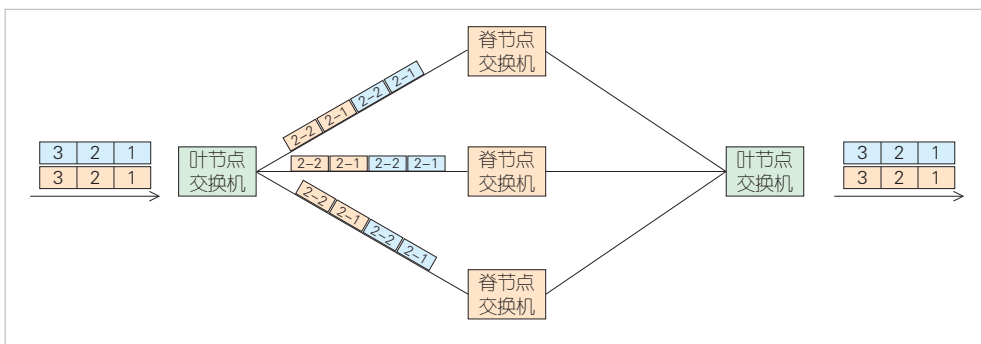
3) 组包交换: 核心思想是将数据包组装成定长且长度较长的数据帧, 并为数据帧添加帧头 (用于组装和还原)。当数据包不足以填充一个大帧时, 就需要填充冗余数据成帧, 并利用网络各节点对大数据帧进行存储转发, 如图 4 所示。该方式下大帧均衡调度的周期短, 可适应高带宽的转发需求。但帧头及冗余数据填充及存储转发机制会带来一定程度的带宽和时延损耗。

基于上述分析, 面向后续智算中心高带宽、低时延的网

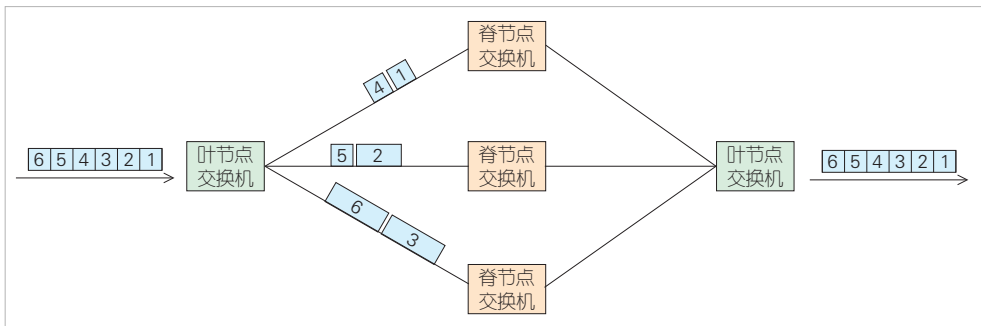
3.2 GSE 技术三大核心理念

3.2.1 基于报文容器的转发及负载分担机制

智算中心网络通常采用胖树 (Fat-Tree) 架构, 任意出入口之间可能存在多条等价转发路径。与云数据中心业务流量不同, 智算业务流量具有“数量少, 单流大”的特点。传统以太网逐流负载分担方式导致链路利用率不均, 从而引起网络拥塞。单流多路径是提升智算中心网络有效带宽、避免网络拥塞的关键技术手段。业界传统网络中实现单流多路径的技术方案包括切包交换、逐包交



▲图2 切包交换示意图



▲图3 逐包交换示意图

络需求，并结合逐包交换方式下即来即转的低时延特性、组包交换方式下的高带宽特性，本文在 GSE 技术架构中提出一种基于报文容器（PKTC）的转发及负载分担机制。该机制根据最终设备或设备出端口，将数据包逻辑分组，并组装成长度较长的“定长”容器进行转发。属于同一个报文容器的数据包被标记为相同的容器标识，沿着相同路径进行转发，以保证同属于一个报文容器的数据包保序传输，如图 5 所示。

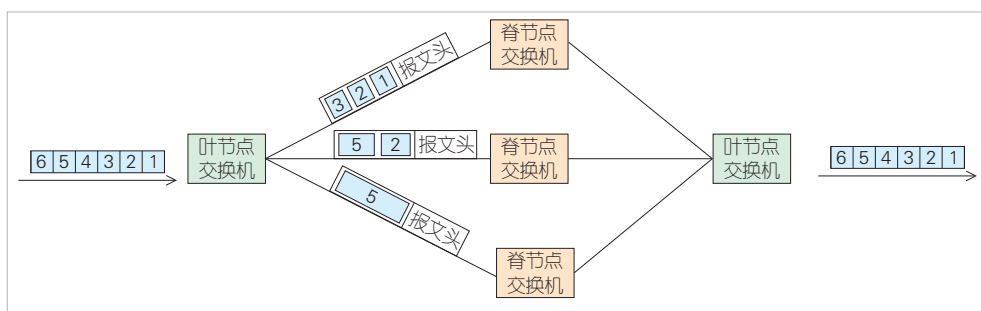
在负载均衡调度时，报文容器被作为转发单位。但由于报文是逻辑组装，无需额外的硬件开销来对数据包进行组装和还原。在网络中转发时添加的报文容器标识，仍以数据包的形式传输，且无冗余数据填充的问题，带宽损耗小。

3.2.2 基于报文容器的 DGSQ 全局调度技术

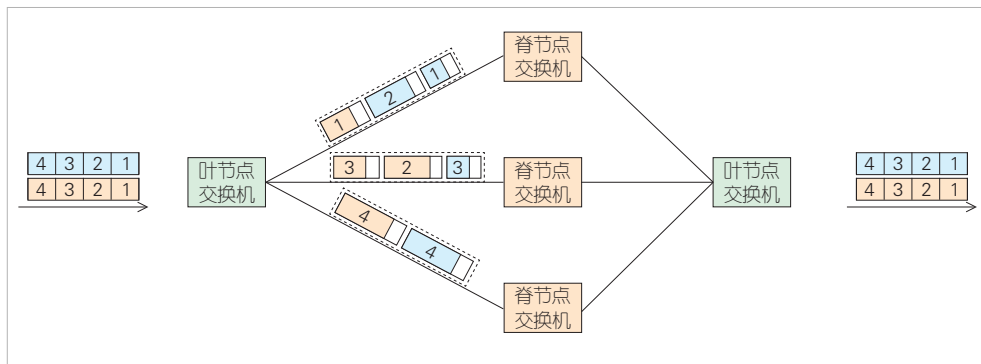
分布式高性能应用的特征是多对一通信的 In-cast 流量模型。如果这种流量是短暂的，在出口处可以通过一定的 Buffer 进行吸收；如果时间持续过长且多个入口的流量相加远大于出口的线速带宽，为了避免丢包，出口设备需启用反压机制保护流量。而反压一旦出现，网络的转发性能就会大幅度下降，从而损害分布式应用的性能。

DCQCN 目前是 RDMA 网络应用最广泛的拥塞控制算法，也是典型的被动拥塞控制算法。发送端根据接收到的拥塞通知报文（CNP）动态调整发送速率。由于 1 个比特的 ECN 信号仅能定性表示网络产生拥塞，但无法定量地表示拥塞程度，所以端侧需要探测式调整发送速率。此外，收敛速度慢会导致网络吞吐性能下降。解决网络拥塞丢包最直接的手段是防止过多的数据注入到网络中造成拥塞，保证网络中任意设备端口缓存或链路容量不会过载。

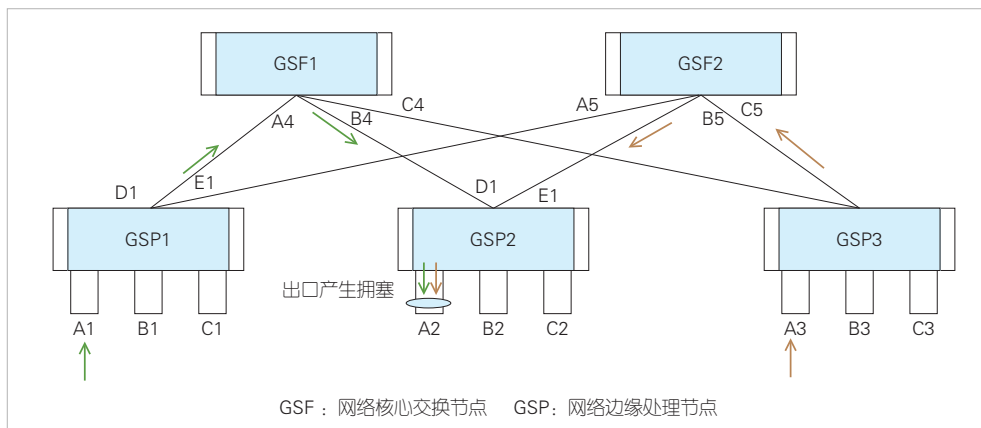
如图 6 所示，GSP1 的 A1 口和 GSP3 的 A3 口同时向 GSP2



▲图 4 组包交换示意图



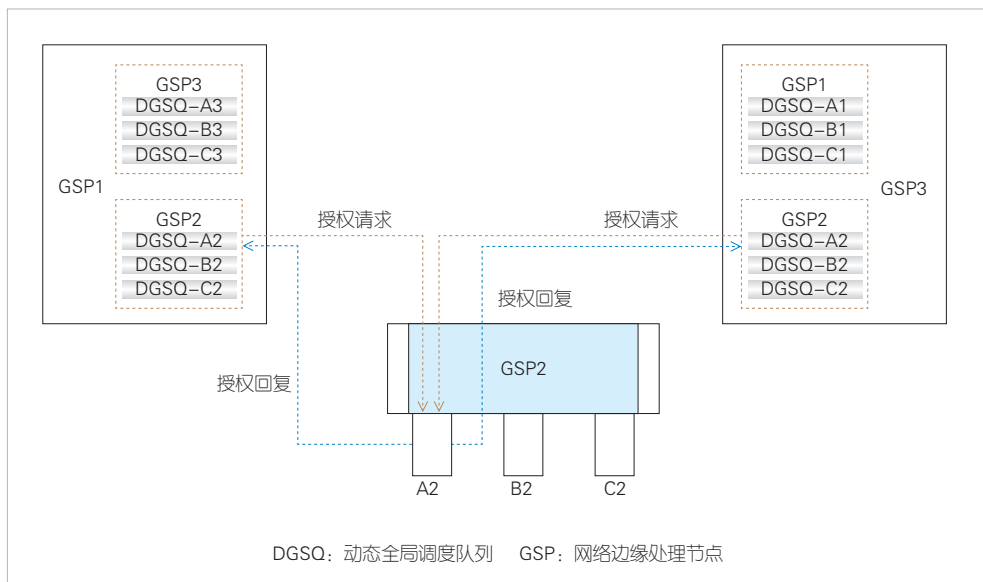
▲图 5 报文容器转发示意图



▲图 6 网络 In-cast 流量发生场景

的 A2 口发送流量，且流量相加大于 A2 的出口带宽。这造成 A2 口出口队列拥塞。这种情况仅通过负载均衡是无法规避的，需要全局控制保证送到 A2 的流量不超过其出口带宽。因此，引入基于全局的转发技术和基于 DGSQ 的调度技术，才可实现全局流量的调度控制。

基于 DGSQ 的全局调度技术如图 7 所示，在 GSP 上建立网络中所有设备出口的虚拟队列，用以实现本 GSP 节点到对应所有出端口的流量调度。本 GSP 节点的 DGSQ 调度带宽依赖授权请求和响应机制，由最终的设备出口、途经的设备统一进行全网端到端授权。由于中间节点的流量压力差异，GSP 去往最终目的端口不再通过等价多路径路由（ECMP）（路径授权权重选择路径，而是需要基于授予的权重在不同



▲图7 基于DGSQ调度流程

的路径上进行流量调度。这种方式可保证全网中前往任何一个端口的流量既不会超过该端口的负载能力，也不会超出中间任一网络节点的转发能力，可降低网络中 In-cast 流量产生的概率，减少全网内部反压机制的产生。

基于 PKTC 的负载均衡技术和 DGSQ 全局调度技术在平稳状态下可很好地进行流量调控与分配。但在微突发、链路故障等异常场景下，短时间内网络还是会产生拥塞，这时仍需要依赖反压机制来抑制源端的流量发送。传统 PFC 或 FC 都是点到点的局部反压技术，一旦触发扩散到整个网络中，会引起头阻 HoL、网络风暴等问题。全调度以太网技术需要精细的反压机制来守护网络的防线，通过最小的反压代价来实现网络的稳定负载。

3.2.3 全调度以太网 GSOS

综合考虑分布式 NOS、集中式 SDN 控制器的优势，全调度以太网 GSOS 分为全调度控制器、设备侧 NOS 两大部分，可全面提升 GSE 网络自动化及可视化能力。

GSP 和 GSF 的盒式设备支持独立部署 NOS，有助于构建出分布式网络操作系统。每台 GSP 和 GSF 具备独立的控制面和管

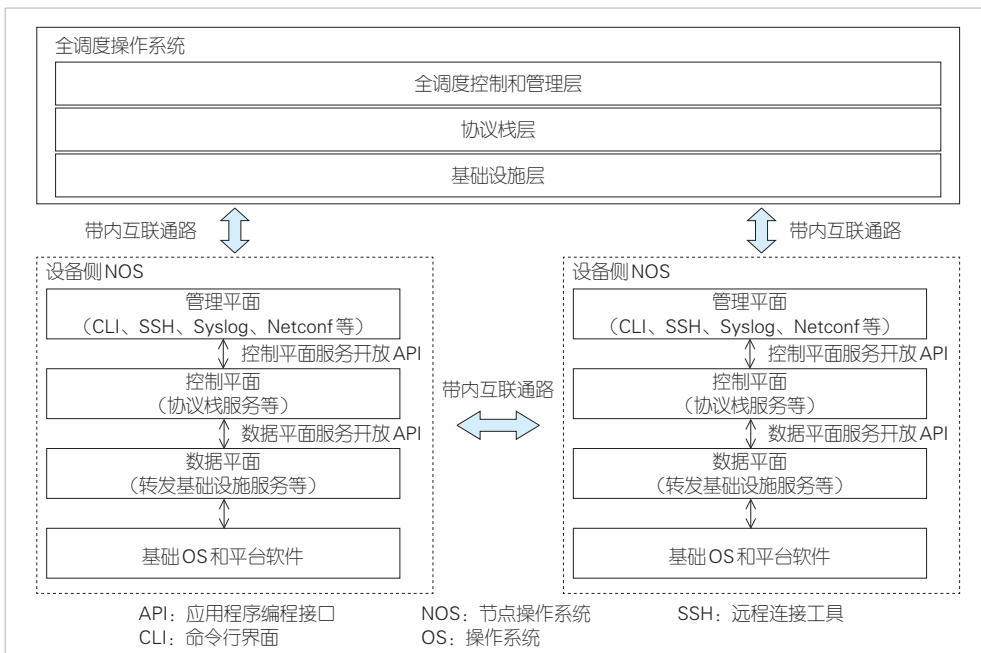
理面，可以运行属于设备自身的网络功能，提升系统可靠性，降低部署难度。分布式 NOS 可以将单点设备故障限制在局部范围，避免对整网造成影响。

集中式 GSOS 提供更好的网络全局信息，简化基于全局端口信息的 DGSQ 系统的建立和维护。同时 GSOS 也是整网运维监控的大脑，可协同设备实现对实时路径、历史的记录及呈现，以支撑网络运维。

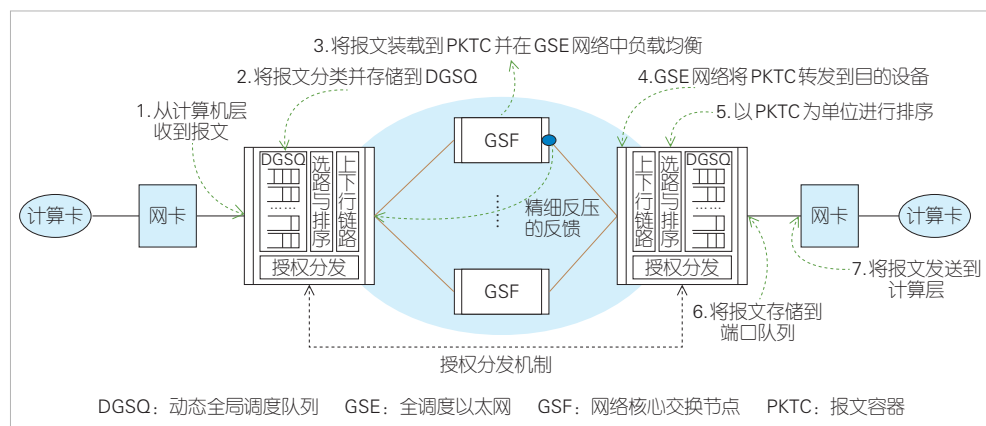
3.3 GSE 网络工作机制

GSE 交换网络采用定长的 PKTC 进行报文转发及动态负载均衡，通过构建基于 PKTC 的 DGSQ 全调度机制、精细的反压机制和无感知自愈机制，实现微突发及故障场景下的精准控制，全面提升网络有效带宽和转发延迟稳定性。相关的具体端到端转发流程图9所示。

- 1) 源端 GSP 设备从计算侧收到报文后，通过转发表找到最终出口，并基于最终出口按需将报文分配到对应的 DGSQ 中进行授权调度。
- 2) 源端 GSP 设备获得授权后，遵循 PKTC 的负载均衡



▲图8 全调度以太网操作系统架构



▲图9 GSE网络端到端流量转发示意图

开发周期长，我们希望各个行业能够携手合作，持续推动相关技术标准发展。

参考文献

- [1] 姚惠娟, 陆璐, 段晓东. 算力感知网络架构与关键技术 [J]. 中兴通讯技术, 2021, 27(3): 7-11. DOI:10.12142/ZTETJ.202103003
- [2] 中国移动通信研究院. 面向AI大模型的智算中心网络演进白皮书 [R]. 2023
- [3] 中国移动通信研究院. 新一代智算中心网络技术白皮书 [R]. 2022
- [4] 中国移动通信研究院. 全调度以太网技术架构白皮书 [R]. 2023

要求，将报文发送到GSE网络中。

3) 当到达目的端GSP设备后，报文先进行PKTC级别的排序，再通过转发表存储到物理端口对应队列，最终通过端口调度发送到计算节点。

作为一种标准开放的新型以太网技术，GSE可采用网卡侧无感知的组网方案，即网络侧采用GSE技术方案，网卡侧仍采用传统RoCE网卡。此外，也可以结合网卡能力演进，将GSE方案各组件的功能在网络组件中重新分工，将部分或全部网络功能下沉到网卡侧来实现。也就是说，在未来的实际应用中，可以将GSP的功能全部下沉到网卡以提供端到端的方案，也可以将网络的起终点分别落在网络设备和网卡上，为后续网络建设和设备选型提供灵活的可选方案。

4 结束语

新型智算中心网络技术已逐渐成为全球创新焦点。智算中心网络是一个多要素融合的复杂系统，是算网的深度融合，它依赖于AI业务、网络设备、交换芯片、网卡、仪表等上下游产业的协同创新。如何提升网络规模和性能，构建超大规模、超高带宽、超低时延的高性能智算中心网络，是提升算力水平的关键。

GSE面向无损、高带宽、超低时延等高性能网络需求业务场景，兼容以太网生态链，通过采用全调度转发机制、基于PKTC的负载均衡技术、基于DGSQ的全调度技术、精细的反压机制、无感知自愈机制、集中管理及分布式控制等技术，实现低时延、无阻塞、高带宽的新型智算中心网络^[4]。该技术架构旨在构建一个标准开放的高性能网络技术体系，助力AIGC等高性能产业快速发展。由于该架构创新难度大、

作者简介



段晓东，中国移动通信有限公司研究院副院长、“新世纪百千万人才工程”国家级人选、教授级高级工程师；长期从事下一代互联网、算力网络、5G网络架构、6G网络架构、SDN/NFV等技术研究工作。



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



王瑞雪，中国移动通信有限公司研究院基础网络技术研究所技术经理、SDN/NFV/AI标准与产业推进委员会（TC610）SDN/NFV技术工作组组长、算网融合产业及标准推进委员会（TC621）国际/开源合作工作组组长；主要研究领域为数据中心网络、SDN/NFV、算力网络等。



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