面向星上大语言模型训练的语义驱动RDMA协议

黄俊凯，潘　恬，杨祎凌，张　娇，黄　韬

（北京邮电大学网络与交换技术全国重点实验室，北京 100876）

摘要：随着大语言模型训练规模持续增长，业界开始探索将部分训练任务部署到轨道空间，依托空间环境中可持续的太阳能供给发展绿色计算，缓解地面训练的能耗压力。然而，多星分布式训练要求星间网络为节点间频繁的同步通信提供高吞吐量、低尾时延以及稳定的传输服务，而星间链路易受环境扰动、丢包现象频发，且星间拓扑随轨道运动周期性动态变化，使传统星间传输协议和地面远程直接内存访问（RDMA）协议均难以直接适配在轨训练场景。为此，提出一种面向星上大语言模型训练的语义驱动RDMA协议——StarRDMA协议。StarRDMA协议充分利用训练通信中各类消息可靠性需求与时效性要求存在显著差异的特征，针对性设计了两大核心机制：一是语义感知的可靠传输服务，协议依据不同训练消息在可靠性需求上的差异，融合选择性纠删编码、选择性确认（SACK）丢包恢复和截止时间感知重传机制，降低高往返时延（RTT）环境下的丢包恢复开销；二是差异化的拥塞控制，结合轨道RTT预测与语义优先级感知的加权加性增大乘性减小（AIMD），使关键路径流量获得更多的带宽分配优势，缓解长RTT流的竞争劣势。基于ns-3仿真平台对StarRDMA协议的性能展开初步评估。仿真结果表明，StarRDMA在高丢包场景下可将AlltoAll、AllReduce集合通信完成时间最高降低72.83%；在后台大流混合共存场景下，StarRDMA差异化拥塞控制可将集合通信完成时间最高降低35.71%。

关键词：太空计算；大语言模型训练；传输层协议；RDMA协议

中图分类号：TP393

文献标志码：A

doi：10.11959/j.issn.2096–8930.tdyt2026028

Semantic-driven RDMA protocol for on-orbit LLM training

Huang Junkai, Pan Tian, Yang Yiling, Zhang Jiao, Huang Tao

State Key Laboratory of Networking and Switching, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: As large language model (LLM) training scales up continuously, there is growing research interest in offloading partial training workloads to orbital space, this paradigm leverages the sustainable solar energy supply in space to develop green computing and alleviate the energy consumption pressure of ground-based training. However, multi-satellite distributed training imposes stringent demands on inter-satellite networks, which must provide high throughput, low tail latency and stable transmission services for frequent synchronization communication among satellite nodes. Nevertheless, inter-satellite links are susceptible to environmental disturbances and prone to packet loss. Meanwhile, the inter-satellite topology changes periodically with orbital motion, rendering both traditional inter-satellite transport protocols and terrestrial RDMA protocols incapable of directly adapting to on-orbit training scenarios. To this end, we propose StarRDMA protocol, a semantic-driven RDMA protocol for on-orbit LLM training. StarRDMA fully exploits the significant differences in reliability and timeliness requirements among various training communication messages, and accordingly designs two core mechanisms. Firstly, it provides a semantic-aware reliable transmission service, according to the differentiated reliability demands of different training messages, the protocol integrates selective erasure coding, SACK-based packet loss recovery, and deadline-aware retransmission, so as to reduce packet loss recovery overhead in high-RTT environments. Secondly, it introduces a differentiated congestion control mechanism. Combining orbit-based RTT prediction with semantic-priority-aware weighted AIMD, the mechanism enables critical-path traffic to gain superior bandwidth allocation, and mitigates the inherent competitive disadvantage of long-RTT flows. This paper presents a preliminary evaluation of StarRDMA's performance using the ns-3 simulation platform. Simulation results

收稿日期：2026–05–06；**修回日期：**2026–08–14

通信作者：潘恬，pan@bupt.edu.cn

基金项目：国家自然科学基金面上项目（No.62372053）

Foundation Item：The National Natural Science Foundation of China（No.62372053）

show that StarRDMA reduces the completion time of AlltoAll and AllReduce collective communications by up to 72.83% under high packet loss scenarios. In the presence of coexist background with large flows, StarRDMA's differentiated congestion control reduces the collective communication completion time by up to 35.71%.

Key words: space computing, LLM training, transport layer protocol, RDMA Protocol.

0　引言

近年来，大语言模型（large language model, LLM）训练正持续向更大参数量、更大数据规模和更大分布式集群规模演进[1-3]。当前大模型的训练往往需要依赖数千乃至数万块图形处理器（GPU）协同完成[4-6]，模型规模扩张带来的不仅是计算需求的激增，也给数据中心的供电、散热与运维保障带来了持续压力。尤其是在训练任务持续数周甚至数月的情况下，巨大的能源消耗正逐渐成为制约大模型进一步扩展的重要因素[7-9]。随着商业航天[14-15]、星间激光通信[16-17]、在轨高性能计算[18]等技术的发展，利用轨道环境承载部分高能耗大模型训练任务，开始从概念走向现实。相较于地面数据中心，在合适轨道条件下可获得持续的太阳能供给，为大规模绿色计算开辟了全新的部署场景与发展边界[10-13]。已有多项工业界与学术界工作尝试在轨部署通用模型能力，并探索空间计算集群部署的可行性[19-22]。

然而，星上大模型训练首先是一个网络问题。无论是“地数天算”，还是直接利用在轨采集数据进行“天数天算”，受限于单星算力与存储资源，大模型训练均难以依靠单颗卫星完成，而必须依赖多星分布式协同计算。与地面GPU集群训练模式类似，星上分布式训练需将整体计算任务拆分并分配至各工作卫星。各卫星节点在执行计算任务的过程中，需与其他卫星节点周期性地进行频繁数据通信以完成模型参数、梯度和激活值的同步更新，涉及AllReduce、AllGather、ReduceScatter等集合通信操作。上述通信操作均位于训练迭代的关键路径，且后续计算任务严格依赖通信同步结果，各卫星节点必须完成全局同步通信后才能进入下一阶段的训练计算，通信完成时间延长将对训练吞吐量、扩展效率和算力利用率产生显著影响。因此，星间网络不再只是单纯的支撑性基础设施，而是直接决定在轨训练效率的核心系统组件。

然而，大模型训练所要求的高吞吐量、低尾时延和稳定完成时间，与星间网络的基本属性存在明显冲突。相较于地面数据中心，星间链路更容易受到姿态扰动、链路对准误差和空间环境变化的影响，因而更易出现误码、短时丢包与带宽波动现象。与此同时，网络拓扑、传播时延与路径状态还会随轨道运动周期性变化。更进一步地，训练流量本身又具有显著的同步突发特征：梯度聚合和参数同步往往在迭代边界集中发生，极易造成网络负载瞬时急剧增加。这种情况下，即使少量的链路扰动也可能演化为全局同步阻塞。

传统星间传输协议，如传输控制协议（transmission control protocol, TCP）、快速UDP互联网连接（quick UDP internet connections, QUIC）和空间通信协议规范—传输协议（space communications protocol specification - transport protocol, SCPS-TP）强调的是复杂链路条件下的通用可靠传输能力，而非大模型训练所需的低尾时延、低抖动、高吞吐量的同步通信能力，其软件协议栈路径较长，中央处理器（CPU）参与度高，难以满足训练通信对低尾时延和开销的严格要求。此外，地面大规模分布式训练广泛采用的远程直接内存访问（remote direct memory access, RDMA）协议在语义和性能目标上与训练需求高度契合，它可通过内核旁路、零拷贝和硬件卸载显著降低主机开销，已成为地面高性能训练集群的重要互联基础。但现有RDMA协议的高性能通常建立在低丢包、低抖动、拓扑相对稳定的网络环境之上，这些假设在星间环境中并不成立。通常来说，直接移植地面RDMA协议既难以应对链路扰动和时变拓扑，也难以适配星上受限的能耗与实现复杂度。这表明星上大模型训练不能简单复用现有星间协议或地面RDMA方案，必须充分考虑星间网络特点与星上资源约束，设计面向在轨训练的专用RDMA传输协议。

为此，本文提出StarRDMA协议，一种面向星上大模型分布式训练的、语义驱动的RDMA传输协议。StarRDMA协议并不试图为所有星间业务提供统一最优解，而是围绕训练通信的两个关键特点进行优化设计。第一，不同训练消息对丢包的容忍度并不相同。某些控制信息、状态元数据和关键参数必须严格可靠传输；某些数值型训练数据则具备一定统计鲁棒性，不必为每一次丢包都付出昂贵的恢复代价；还有一些小而关键的时延敏感消息，更适合通过一次发送中的冗余保护来避免丢包后往返时延（round-trip time, RTT）级的重传等待。第二，不同训练流量对时效性的要求也不相同。某些流

量直接位于训练关键路径上，而另一些后台流量即便有延迟，对当前训练性能的影响也有限。如果传输层仍使所有流量公平共享稀缺的带宽资源，将会在网络负载高峰时延长关键路径的通信完成时间。

基于上述观察，StarRDMA 协议主要包含两大核心机制。其一是语义感知的传输可靠性机制。StarRDMA 协议提供严格可靠、部分可靠和时延优先 3 类服务，由通信库负责根据训练消息对传输可靠性的敏感程度将其映射至对应的可靠性等级。对于时延敏感短消息，StarRDMA 协议引入选择性纠删编码，以提高一次传输成功率、降低对端到端重传的依赖；对于需要恢复的丢包场景，则采用基于选择性确认（selective acknowledgment, SACK）的精细化反馈机制提升重传效率，同时设计截止时间感知的重传调度策略，使有限带宽优先服务对训练进度有重要价值的消息。其二为差异化服务的拥塞控制机制。针对星间网络中显著的 RTT 异构特性，StarRDMA 协议结合轨道几何和路由先验信息预测基础 RTT，并在加性增大乘性减小（additive increase multiplicative decrease, AIMD）拥塞控制框架中引入加权补偿策略，以减弱长 RTT 流在带宽竞争中的天然劣势。与此同时，StarRDMA 协议将应用暴露出的语义优先级映射为差异化增速权重，使关键路径流量在共享瓶颈链路时能够获得带宽分配优势。

本文在 ns-3 仿真平台上实现 StarRDMA 协议，并基于 24×24 2D-mesh 动态周期运行拓扑进行初步的仿真评估和性能分析。仿真结果表明，StarRDMA 协议在高丢包场景下可将 AlltoAll、AllReduce 集合通信完成时间最高降低 72.83%，优于 RoCEv2 协议；在后台大流混合共存场景下，StarRDMA 协议差异化拥塞控制可将关键路径集合通信的完成时间最高降低 35.71%。

1　背景与挑战

1.1　星上大模型训练通信需求

1.1.1　大模型训练上星的动因

近年来，大模型训练正持续朝着更大参数量、更大训练数据规模和更大集群规模的方向演进[1-3]。当前前沿模型的训练任务通常依赖数千乃至数万块 GPU 协同完成[4-6]。随着训练规模不断扩大，模型训练不仅对计算资源提出了极高要求，也显著加剧了地面数据中心在供电、散热以及运维保障上的压力。尤其在单次训练任务持续数周甚至数月的背景下，大模型训练带来的高额能源消耗问题，正逐步成为制约其进一步发展的关键因素之一[7-9]。

在此背景下，业界与学术界开始重新思考计算基础设施的部署边界，并探索利用轨道环境承载部分高能耗训练任务的可能性。与地面数据中心相比，轨道空间具备持续、稳定且高强度的太阳能供给条件，从理论上为大规模绿色计算提供了新的承载平台[10-13]。与此同时，随着商业航天产业的快速发展，卫星发射成本持续降低[14-15]、星间激光通信技术逐步成熟[16-17]，加之空间高性能计算平台的迭代升级[18]，卫星正从单纯的数据采集与转发节点逐步演进为可执行复杂智能计算任务的在轨计算节点[10,13]。在此基础上，星上智能计算开始从概念探索走向原型验证。国星宇航[19]、Starcloud[20]、之江实验室[21]等机构已将通用大模型部署至卫星平台，开展星上大模型推理乃至训练相关能力的概念验证；Google 推出的“捕日者计划”尝试通过构建太空计算集群突破地面数据中心在能源与冷却方面的瓶颈[12]；SpaceX 也提出了发射百万级卫星、构建绿色轨道数据中心的设想[22]。太空数据中心进行星上大模型训练的系统概念如图 1 所示，其中 ISL 指星间链路，DSA 指特定领域架构。

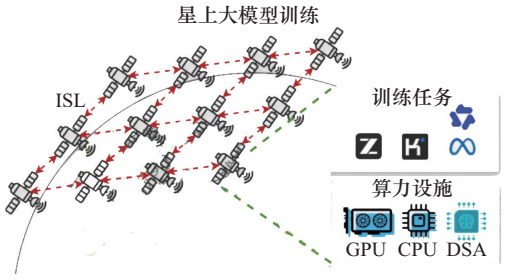
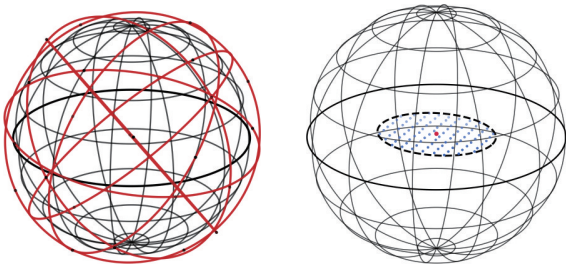


图 1　星上大模型训练的系统概念

从任务类型看，星上训练模式大致可分为两类：一是地数天算，即训练数据主要由地面产生或在地面侧汇聚，再上注至轨道环境，由轨道计算设施承担训练任务；二是天数天算，即卫星直接利用在轨采集的数据开展数据处理、增量学习乃至模型训练，仅按需回传结果、摘要信息或模型参数。前者强调利用轨道能源与在轨算力分担地面训练压力，后者强调利用在轨数据的实时性、局部性与低回传开销实现“采集—处理—更新”的闭环。无论采用哪种模式，受限于单星算力资源约束，当模型参数量及数据规模较大时，系统均需依托多星分布式协同计算完成训练。各卫星节点在执行计算任务的同时，需要频繁通过星间网络进行模型参数、梯度以及激活值的同步更新，训练任务的推进依赖节点间的同步通信结果，因此星间高性能通信是星上分布式训练亟须重点考虑的核心问题。

需要指出的是，星上大模型训练不能被简单理解为将现有天基推理或星上数据处理工作的规模放大。原因是：现有天基计算大多围绕在轨推理、遥感处理或任务卸载展开，其负载通常依赖单星或少量卫星节点的计算能力，跨节点耦合相对较弱，而大模型训练需要大量的卫星节点协同进行，本质上是一个长时间运行、强同步的分布式过程，对稳定、高吞吐量、低时延的星间互联通信具有更强的依赖性；另外，在卫星拓扑上，训练型轨道计算设施并非一定要依赖Walker星座这类全球覆盖星座架构，不同于推理型系统的全球覆盖需求，训练型系统更关注节点之间能否长期维持稳定、高性能的互联网络，以支撑频繁的模型同步更新，因此，其更可能采用局部高密度部署、紧凑编队飞行的组织形式。太空数据中心卫星拓扑（对比 Walker 星座）如图 2 所示。以Google提出的多星紧凑编队设想为例（图2（b）），其星间距离可缩短至几千米甚至数百米，与传统全球星座中数百千米量级的星间距离形成鲜明对比，也为大模型训练所需的低时延互联提供了可能[12]。



(a) Walker星座（24/6/1）　　(b) 紧凑编队星座

图2　太空数据中心卫星拓扑（对比Walker星座）

1.1.2　分布式训练通信模式与需求

多星分布式协同训练首先需要将训练任务划分给各颗计算卫星。类似于地面GPU训练集群，星上训练可以采用数据并行、张量并行、流水线并行等策略对训练任务进行划分，这些并行方式都会引入频繁的跨节点通信，以在每次迭代时完成模型参数、梯度等数据的聚合同步。例如，数据并行中的AllReduce集合通信用于聚合各节点本地计算得到的梯度，以保证参数更新的一致性；张量并行中的AllGather、ReduceScatter集合通信用于在不同节点之间交换、重构以及归约分片的激活值、参数或梯度；流水线并行中的点到点传输则用于在相邻阶段之间传递前向激活和反向梯度。这些通信操作并不是训练过程之外的附属开销，而是直接位于训练迭代的关键路径上。例如，若某个节点的梯度聚合未完成，参数更新便无法继续；若激活或梯度未及时到达，后续流水阶段便无法推进。因此，对于大规模分布式训练而言，网络不再只是支撑性基础设施，而是直接决定训练吞吐量、扩展效率和算力利用率的关键组件。因此，星上训练对于高性能、低时延、稳定的星间互联网络有着迫切需求。

然而，上述需求与卫星网络的基本属性之间存在明显冲突。训练希望获得接近地面高性能集群的稳定高吞吐量、低尾时延和低开销通信能力，然而星间链路却更容易受到环境扰动影响，且网络拓扑和路径状态会随轨道运动发生变化，这种冲突给星间高性能传输协议的设计带来了巨大挑战。

1.2　RDMA协议基础与星地协议对比

1.2.1　地面RDMA协议

RDMA协议是一类面向高性能计算与存储互联的远程直接内存访问技术，被广泛应用于GPU集群的高性能互联。RDMA允许主机直接访问远端主机的内存，在数据传输阶段无须远端主机CPU与操作系统参与数据拷贝。RDMA协议的典型实现包括InfiniBand、RoCEv2和iWARP[23-24]。与TCP、QUIC等通用传输协议不同，RDMA协议并不以兼容广域互联网和通用字节流传输为首要目标，而是围绕低开销、高吞吐量、低时延和高稳定性的集群通信需求进行设计。其通过内核旁路、零拷贝以及传输协议硬件卸载，显著减少CPU干预、协议处理与内存搬运开销，从而在高带宽环境下仍能维持较高链路利用率和较低尾时延。RDMA协议的上述能力对于大模型训练而言尤为重要，借助RDMA协议服务，训练节点可以高效地完成梯度聚合以及参数同步，提高GPU训练吞吐量。RDMA协议也因此成为地面训练集群的主流互联技术。

1.2.2　现有卫星传输层协议与地面RDMA协议对比

当前星地及星间网络中常见的传输层方案主要包括TCP、QUIC以及面向空间环境优化的SCPS-TP等。此类协议大多继承了通用网络传输的设计思路，重点面向复杂链路条件下的端到端可靠传输、公平共享和广域互联兼容性。上述3种传输层协议与地面RDMA协议在关键机制上的差异对比见表1。

RDMA协议的显著优势在于内核旁路、零拷贝与协议的硬件卸载，这使得在数据路径上CPU或GPU只需提交通信请求，而不需要在用户态和内核态之间进行数据拷贝，也不需要承担传输协议的处理，而是全部交由RDMA网卡完成，从而释放了CPU或GPU的计算资源。相比之下，在典型实现中，TCP、QUIC以及SCPS-TP协

表 1　TCP、QUIC、SCPS-TP 与地面 RDMA 协议的关键机制对比

对比维度	TCP	QUIC	SCPS-TP	RDMA 协议
传输语义	面向有序字节流，应用按流读写数据	基于 UDP 提供多路字节流，也支持数据报	延续 TCP 风格的字节流语义	面向消息/远程内存访问语义
顺序保证	按字节流顺序交付	每条流内顺序交付，不同流相互独立	按字节流顺序交付	按照消息提交顺序交付
CPU 开销	软件协议栈，报文分段、拥塞控制、丢包检测等均依赖 CPU	软件协议栈，用户态加解密、丢包检测、拥塞控制均消耗 CPU	协议处理主要仍由主机侧 CPU 承担	CPU 负责投递工作队列元素、轮询完成队列；协议处理、数据搬运均由网卡完成
内核旁路	传统内核协议栈仍依赖内核处理	部分用户态化，但底层仍通过 UDP socket 经过内核 I/O 路径	通常不支持旁路内核	支持，用户进程可直接与 RDMA 网卡队列交互
数据拷贝	通常存在用户态/内核态拷贝，接收端还需从内核拷贝回应用	虽在用户态实现部分逻辑，但数据收发仍绕不开 socket 缓冲路径	一般不具备直接数据放置能力；需 CPU 拷贝	基于内存注册+直接内存访问，支持 CPU 零拷贝；RDMA Write 可直接写远端内存
可靠性机制	基于序列号+ACK/SACK +快速重传 + 重传超时	基于序列号+ ACK + 探测超时	在 TCP 式确认/重传框架上做空间增强	可靠模式下由 RDMA 网卡维护报文序列号、ACK/NACK、重传队列
拥塞控制机制	常用 Reno/CUBIC/BBR，依据 ACK、丢包、RTT 调整拥塞窗口	通常沿用 TCP 类拥塞控制，在用户态执行	多基于窗口/确认机制做长时延链路适配	InfiniBand 常用信用式流控；RoCEv2 常结合显示拥塞通知机制+DCQCN，端网协同控速
对网络环境的要求	可适应广域、异构、动态网络，但代价是协议路径较长	适应复杂互联网环境，兼顾网络地址转换穿越与安全传输	面向高时延、高误码空间链路优化	高性能通常依赖低丢包、低抖动、拓扑较稳定以及专用硬件网卡支持
对训练通信的适配性	通用可靠传输能力强，但软件栈开销和尾时延偏大	灵活性高，但用户态软件处理和加密开销不利于高频模型参数同步	更适合传统空间数据传输，不适合高频强同步训练通信	高性能、低时延，适合高频强同步训练通信

议均为软件协议栈，依赖通用 CPU 进行协议处理和数据搬运。RDMA 协议和内核 TCP 数据路径对比如图 3 所示，以内核 TCP 协议栈为例，数据发送通常需要应用程序将数据写入用户态缓冲区，经系统调用陷入内核后再复制到内核发送缓冲区，随后由内核协议栈完成分段、拥塞控制、重传管理和报文发送，接收端还需经过内核收包、协议处理以及数据从内核态拷贝回用户态等过程。由于协议完全运行于软件协议栈，因此一方面 CPU 资源占用较高；另一方面软件栈调度开销大、抖动与稳定性较差，无法满足高吞吐量、低时延的传输性能要求。相比之下，在 RDMA 协议中，发送端应用只需准备好通信请求并提交给 RDMA 网卡，由 RDMA 网卡完成后续的协议处理和数据搬运，将数据直接写入对端注册内存即可。在上述过程中，RDMA 协议的写/读操作甚至无须远端 CPU 的参与，消除了数据路径上繁重的软件栈开销和 CPU 资源占用，实现了更优的性能表现。

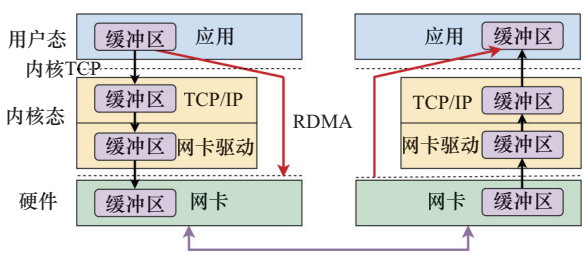


图 3　RDMA 协议和内核 TCP 数据路径对比

在设计目标上，TCP、QUIC 和 SCPS-TP 更强调在动态、异构甚至高误码环境下的通用可靠传输能力，而 RDMA 协议的高性能通常建立在网络负载可控、拓扑相对稳定、端侧具备专用硬件支持等条件之上。也就是说，RDMA 协议的设计理念虽然与大模型训练需求高度契合，但其既有实现往往依赖地面数据中心或高性能集群环境，难以直接迁移到星间网络。这也说明，星上大模型训练所需的传输协议既不能简单沿用现有的卫星传输机制，也难以直接使用地面 RDMA 协议，而需要针对星间环境特性重新设计。

1.3　星上 RDMA 协议的设计挑战

尽管 RDMA 协议在大模型训练通信中表现出显著优势，但在复杂的星间网络环境中，构建一个高性能的星上 RDMA 协议仍面临诸多挑战。

1) 星间链路状态较地面环境更不稳定、带宽资源更紧缺，难以满足地面 RDMA 协议所依赖的无损网络假设。地面 RDMA 协议通常建立在低丢包、低抖动和高稳定性的网络环境中，并借助链路层机制尽可能维持无损传输，以避免丢包对传输性能的影响。然而，星间链路更容易受到姿态扰动、链路对准误差、空间环境变化等因素影响，链路误码和短时丢包更难避免；同时，较长的传播距离也使得一次丢包或重传带来的时延代价明显高于地面数据中心。因此，地面 RDMA 主流以太网部署方案依托基于优先级的流量控制（Priority-based flow control，

PFC）实现的无损传输难以在星上环境实现。另外，星间链路带宽资源稀缺，地面RDMA协议所依赖的Go-back-N重传机制带来的带宽开销较大，难以应对高丢包率场景下高效重传的需求。

2) 周期性拓扑变化和链路可用带宽波动使拥塞控制更难保持稳定。与地面数据中心相对稳定的网络拓扑不同，星间网络的邻接关系、传播时延和可用带宽会随轨道运动发生周期性变化。然而，链路状态变化未必由真实拥塞引起，却可能被传统拥塞控制机制误判为拥塞信号，导致不必要的速率收缩或利用率下降。对于依赖稳定时延反馈和相对固定路径条件的地面RDMA协议拥塞控制方案而言，这种时变特性会显著增加控制难度。

3) 大模型训练流量具有明显的同步突发特征，会进一步放大上述问题。在分布式训练过程中，梯度聚合、参数同步等集合通信通常在迭代边界集中发生，易形成高并发、强同步的突发流量。当这种训练流量与时变拓扑、链路波动叠加时，网络更容易出现瞬时排队积压以及尾时延放大，从而使少量链路波动演化为全局同步阻塞。

4) 星上算力、散热与能耗约束要求协议复杂度必须可控。地面高性能RDMA协议的实现通常依赖功能较强的RDMA网卡、充足的缓存资源以及较宽松的功耗和散热条件，而这些前提在星上平台往往受到显著限制。协议在可靠性维护、拥塞控制或状态管理上过于复杂，不仅会增加实现成本，还可能额外占用有限的计算、存储与能耗预算，进而挤占本该应用于模型训练的系统资源。因此，面向星上训练的高性能传输协议也应尽可能保持机制简洁、状态可控和实现可部署。

综上，星上大模型训练所需的传输协议既不能简单沿用现有的卫星传输协议，也不宜直接使用地面RDMA方案，其设计需要同时考虑卫星网络的特点以及星上资源约束情况。

2　StarRDMA协议设计

2.1　协议概述

前述分析表明，星上大模型训练所面临的核心通信矛盾包括星间链路带宽有限、链路不稳定易丢包，以及星间距离较长导致的不可忽略的传播时延。这些问题使得一个在地面数据中心表现良好的通用型高性能传输协议在星上环境中往往难以同时兼顾吞吐量、时延和稳定性。本文并不试图为所有星间业务设计一种普适最优的RDMA协议，而是根据大模型训练的通信特点，设计一种面向星上大模型训练的专用传输协议——StarRDMA协议。

StarRDMA协议的设计基于两个直接的观察：一是不同训练消息或参数块具有差异化丢包容忍度，无须对所有类型的训练流量均提供严苛的可靠传输保障，可依据训练业务语义按需适配差异化可靠服务等级；二是不同通信流量的紧迫性也不相同，某些通信流量并不会直接进入训练流水线的主路径，若传输层采用统一的带宽分配策略，就会在关键流量最需要资源时把带宽平均摊薄，导致训练效率显著下降。

图4展示了StarRDMA协议的整体架构，协议主要包含以下两个关键机制。① 语义感知的传输可靠性机制。StarRDMA协议提供多种可靠性服务，由应用或通信库根据模型参数对丢包的敏感程度将训练消息映射到不同的

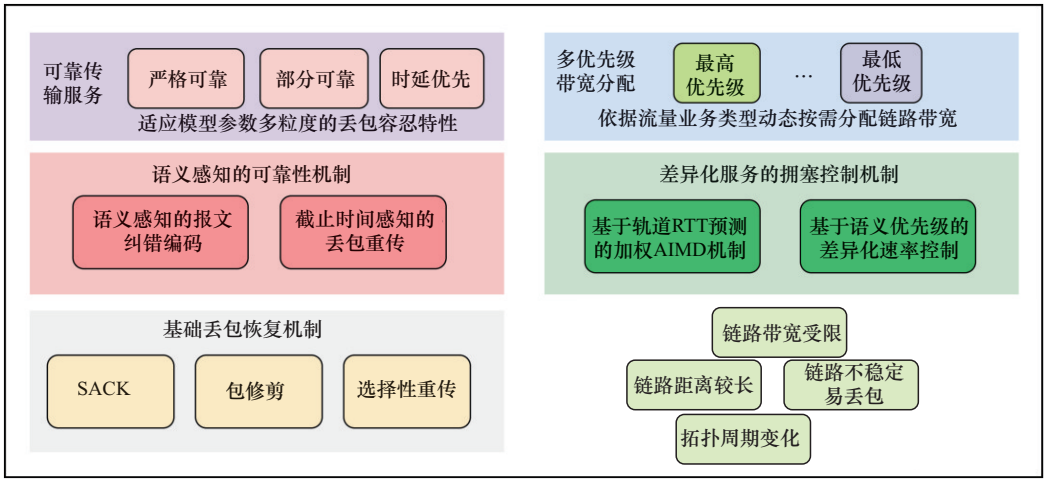


图4　StarRDMA协议的整体架构

可靠性等级。对于必须完整到达的数据，StarRDMA协议提供严格可靠传输服务；对于具备丢失容忍性的流量，StarRDMA协议允许丢包时不完全恢复；对于难以承受重传开销的时延敏感短消息，StarRDMA协议则进一步结合冗余编码，在不等待往返重传的情况下提升一次传输成功率。② 差异化服务的拥塞控制机制。考虑到星间网络是一个具有显著RTT异构性的多跳网络，AIMD类拥塞控制在这种环境下容易因反馈时延不同而产生带宽不公平的现象。StarRDMA协议因此引入基于轨道RTT预测的加权AIMD机制，将已知的轨道几何与路由信息纳入速率调整过程，以削弱RTT差异的影响。同时，StarRDMA协议进一步根据应用暴露出的流量优先级为不同类型的流量分配不同的增速权重，使关键路径流量在竞争有限链路带宽时能够获得更多的带宽。StarRDMA协议基于跨层协同的思想进行设计，使传输层同时具备网络状态感知与训练任务感知的双重能力。

2.2 语义感知的传输可靠性机制

2.2.1 模型参数的多粒度丢包容忍特性

在地面GPU集群中，互联网络大多默认建立在高带宽、低时延、低误码率的前提下，因此端到端协议通常倾向于为所有训练报文提供统一且强一致的可靠语义，然而这一做法在星上环境中并不适用。原因是：除拥塞导致的队列溢出之外，由链路误码引发的报文损坏或丢弃同样是一个不可忽视的因素；当链路带宽本身有限、往返时延又较高时，频繁依赖端到端重传来保证传输的完整性，往往会使大量带宽被消耗在丢包恢复的流量上，造成有效吞吐量的下降，进而拖慢集合通信完成时间。一个关键的机会在于：模型训练中的不同参数与中间数据对于丢包的敏感程度并不相同。从训练语义上看，这种差异可以体现在多个粒度上。① 不同消息类型的容错性不同。例如优化器状态、参数分片元数据等消息通常属于强语义一致性数据，这类消息一旦丢失或损坏，不仅造成简单的数值噪声，而是可能会导致训练状态不一致，甚至直接触发训练失败。对于这类消息，协议必须提供严格可靠性。与之相对的是，一些数值型训练数据具备一定的统计鲁棒性。以梯度同步为例，深度学习优化本身就对噪声具有一定的容忍能力，小比例梯度元素的延迟或缺失，在许多优化设置下并不会立即造成收敛失败，只会在一定程度上影响收敛速度或最终精度。例如，混合专家模型训练中的Token路由结果，若个别小报文因误码丢失，在某些实现中可通过局部丢弃、回退路由等方式吸收其影响。也就是说，这类消息并不一定需要追求严格的可靠性。② 在同一类消息内部，不同参数块的重要性也可能不同。在大模型训练中，参数和梯度往往以块状的形式传输（例如bucket、chunk），而这些块状数据对训练效果的影响并非完全均匀。例如，在量化或稀疏训练中，被保留的Top-K元素天然比被截断的长尾元素更重要，这意味着协议层可以进一步利用来自框架侧的统计信息，在块粒度上实施差异化可靠性。

基于上述分析，StarRDMA协议将可靠性服务划分为3类，并允许应用或通信库在提交消息发送请求时显式声明以下服务。

1) 严格可靠服务：适用于绝不允许丢失或损坏的消息，协议保证消息完整、有序地交付给上层。

2) 部分可靠服务：适用于允许有限范围内丢失，但要求丢失受控的消息，对于此类数据，协议并不承诺恢复每一个丢失分组，而是允许在满足预算约束的前提下跳过部分丢失分组恢复。

3) 时延优先服务：适用于对时延敏感的小消息，对于此类消息，协议优先追求一次传输成功率和低尾时延，而不是丢包后重传恢复。

2.2.2 语义感知的数据报文编码机制

在星上环境中，仅靠端到端重传来提供可靠性的代价往往过高。原因是：即便采用紧凑编队飞行来缩短星间距离，星间传播时延仍明显高于地面机架内环境，因此一旦发生丢包，等待反馈并完成重传会显著拉长消息完成时间。这一问题对短小且频繁的关键路径训练流量影响尤为突出。此类消息一旦发生丢包，恢复过程通常会引入至少一个RTT量级的额外等待；若进一步出现连续丢包或重传报文再次丢失，恢复时延还可能累积至多个RTT，从而显著拉长消息完成时间并阻塞后续计算。

基于这一观察，StarRDMA协议在传输层引入选择性的报文纠删码（erasure coding, EC）。其目标不是对所有流量统一增强可靠性，而是仅对那些重传代价高于冗余代价的消息提供一次发送内的冗余保护，从而降低消息级恢复对额外RTT的依赖。具体地，发送端将一条消息划分为k个数据报文，并生成r个冗余报文，形成一个$(k+r, k)$编码组；接收端一旦收到任意k个报文，即可恢复原始消息。采用系统化Reed-Solomon码作为默认编码，因为它能够在较小编码组下提供确定性的擦除恢复能力，并且可在无丢包时直接交付原始报文，不引入额外解码路径。

StarRDMA协议不对所有消息启用纠删码，持续性的冗余会挤占本就稀缺的星间带宽，同时编码和解码本身

也会带来额外开销。因此，StarRDMA协议仅对时延优先的流量启用编码，对于其他类型的流量，仍以显式确认与重传作为主要恢复手段，因为这类消息通常为大流量，为大流量持续引入冗余会带来大量的带宽开销。

2.2.3　截止时间感知的丢包重传机制

虽然纠删编码可以减少部分重传，但它无法彻底避免丢包重传，尤其对严格可靠服务来说，显式重传仍然是必要的。问题在于，在星间链路带宽受限的条件下，重传流量本身会与正常训练流量竞争带宽。如果重传策略盲目地对所有消息执行重传，而不考虑消息对丢包的容忍程度以及消息的时效性，就可能出现这样一种低效的情况：网络正在消耗链路带宽去恢复一批对丢包有较强容忍的消息或是即将失去价值的旧消息，而真正对训练效果有关键影响的消息却迟迟得不到充足的恢复。为避免这一问题，StarRDMA协议设计了一个截止时间感知的丢包重传机制，由应用或通信库根据消息的丢包容忍程度设置一个截止时间，协议负责根据截止时间确定重传的先后顺序。

2.2.3.1　基础丢包重传机制

StarRDMA协议的基础丢包恢复机制建立在SACK之上，借鉴TCP SACK的思想，由接收端精细地告知发送端哪些报文仍未收到，使得发送端可以精准确定丢失需重传的报文，避免盲目地进行Go-back-N重传导致链路带宽浪费。

1) 基于位图的接收状态记录与SACK反馈。在每个RDMA队列对（queue pair, QP）下，接收端维护一个按报文序列号（packet sequence number, PSN）组织的位图，用于标记当前已收到的数据报文。与累计确认不同，位图能够精确表示消息中的空洞位置。接收端在连续收到n个报文都仍未将累计确认PSN向前推进时，触发发送SACK，SACK携带接收位图返回给发送端，发送端根据SACK中的空洞位置构造重传请求，并在拥塞控制允许的情况下重发丢失的报文。

2) 面向拥塞丢包的包修剪（packet trimming）机制。packet trimming是地面GPU数据中心高性能交换机上的一项新兴丢包通知机制[25]。当网络节点检测到即将发生队列溢出时，传统方法通常直接丢弃整个报文，而packet trimming则是将溢出的报文负载截断，仅保留精简后的报文头部并转发给接收端。接收端一旦收到这类截断报文，即可获知对应报文已在途中遭遇拥塞丢弃，从而无须等待达到累计阈值便主动触发SACK，并在其中明确携带被丢弃的序列号，显式触发报文重传。

在有损且易丢包的卫星网络环境中，相较于地面数据中心网络，流尾包丢失的影响更为显著。流尾包丢失后，由于发送端已无后续报文发送，接收端无法通过后续报文到达来判断是否存在尾部空洞，只能依赖发送端基于重传超时（retransmission timeout, RTO）的超时重传机制发现该丢包事件。这会显著拉长流完成时间，特别是会对短流造成严重影响。为此，StarRDMA协议在交换机侧引入packet trimming机制，重点防护尾部报文丢失。具体而言，交换机根据RDMA报文基础传输头部（base transport header, BTH）中的Opcode字段值是否为Last或Only来判断当前报文是否为一条消息的尾部报文。这里并不直接识别流尾部，是因为在缺少应用层提示的前提下，协议层无法判断应用后续是否仍会提交新的消息发送。因此，识别消息尾部是一种更容易实现的折中方案。当消息尾部报文因队列拥塞即将被丢弃时，交换机采用截断转发的方式，将仅包含关键头部信息的尾部报文继续交付给接收端，从而显式触发接收端发送SACK反馈，避免尾包丢失只能依赖RTO来恢复。

此外，由于SACK报文需要携带接收位图，其报文长度和带宽开销相对较高。为避免频繁发送SACK带来的额外开销，StarRDMA协议默认仅在累计确认PSN连续n个报文未被推进时才触发SACK。然而，对于小消息而言，该策略可能在丢包时引入额外的尾部等待时延，因为小消息本身包含的报文数量有限，可能难以及时满足连续n个报文未推进的触发条件。因此，StarRDMA协议也为小消息开启了packet trimming机制。StarRDMA协议通过为大、小消息设置不同的差分服务代码点（DSCP）值来标识小消息，但仍使其共享相同的RDMA网卡交换机物理队列。发送端RDMA网卡在收到应用提交的工作队列元素（work queue element, WQE）后，根据WQE中提供的待发送数据大小是否低于预设的小消息边界阈值，判断该消息是否属于小消息，并据此设置对应的DSCP值。交换机随后可基于该DSCP标记对小消息报文启用packet trimming，显式反馈小消息丢包。

2.2.3.2　截止时间感知的重传队列调度

通过上述基础的丢包恢复机制，StarRDMA协议可以快速检测到丢包重传需求。但在星间链路带宽很有限的情况下，重传恢复对链路带宽的消耗也应被谨慎考虑，带宽资源应被优先用于恢复对训练效果有关键影响的消息，以保证其在丢包时能够尽早恢复。对于同一类可靠性服务等级的消息来说，对丢包容忍度越低的消息的恢复紧迫性更为强烈。基于这一考虑，StarRDMA协议为每

条消息维护一个消息级传输截止时间D_m，用以衡量消息对丢包的容忍程度。

消息级截止时间D_m由上层训练运行时给出，并通过WQE传递给发送端RDMA网卡。StarRDMA协议并不要求网络协议层直接理解训练语义，而是将截止时间作为应用层提供的调度提示。对于分布式训练任务，通信操作通常嵌入确定性的迭代执行图中，其输出张量会在后续计算、同步或参数更新阶段被消费。因此，训练运行时可以根据当前通信操作的发起时间、后续依赖计算的预计执行时间，以及该消息被下游算子首次消费的预计时间，估计该消息的最晚完成时刻。具体地，若记消息m的发送时间为t_m^{send}，其结果被下游算子首次消费的预计时间为t_m^{use}，并为运行时抖动、链路RTT波动以及拥塞排队时延等开销预留一定的安全裕量Δt_{margin}，则可将消息级截止时间设置表示为式(1)：

$$D_m = t_m^{\text{use}} - \Delta t_{\text{margin}} \tag{1}$$

其中，t_m^{use}可以由训练框架根据静态执行图、历史迭代profiling结果或在线滑动平均估计得到，Δt_{margin}则用于避免估计误差导致的恢复调度过于激进。对于难以精确获得下游消费时间的场景，应用也可以按照消息类型或训练阶段提供默认截止时间。StarRDMA协议将该截止时间作为重传调度提示使用，并不保证消息一定在截止时间前完成，而是在重传队列存在竞争时决定有限链路带宽应优先恢复哪些消息。

基于这一截止时间提示，发送端在维护重传队列时不再单纯按报文丢失的先后进行排序，而是为每个待恢复报文计算其所属消息的剩余紧迫度，如式(2)所示：

$$\text{urgency}(\text{pkt}) = D_m - t_{\text{now}} \tag{2}$$

紧迫度越高意味着：若不立即恢复，该消息越有可能在截止时间前无法完成。

在此基础上，StarRDMA协议采用分层的截止时间感知重传调度机制：① 按照可靠性服务等级排序（时延优先服务>严格可靠服务>部分可靠服务），服务等级之间采用加权轮询的方式为不同服务等级分配重传带宽，以避免低优先级服务长期饥饿；② 在同一服务等级内部，采用最早截止时间优先（earliest deadline first, EDF）调度策略，优先恢复截止时间更近、剩余紧迫度更高的消息。

对于部分可靠服务，当消息截止时间到达时，若对应数据的重传仍未完成，StarRDMA协议将主动终止该消息的后续重传，由应用自行判断当前已传输的数据是否满足业务需求，若无法满足业务需求则由应用重新发起消息传输。为防止未恢复报文造成的PSN空洞阻塞后续消息，发送端发送Skip-PSN报文，显式指示接收端跳过当前消息的剩余缺失报文，并将期望PSN推进至下一条消息的起始PSN。该机制在保留部分可靠语义的同时，避免了被放弃消息对后续可靠传输产生持续阻塞。

2.3 差异化服务的拥塞控制机制

2.3.1 基于轨道RTT预测的加权AIMD

与地面数据中心常见的Fat-tree、Spine-leaf或Dragonfly等扁平拓扑相比，星间网络拓扑通常更接近规则的2D Mesh拓扑。任意两个训练节点之间的通信很少通过单跳完成（往往需要经过多个中继节点完成），同时，由于卫星之间的物理距离远大于地面数据中心机架内交换芯片之间的距离，因此不同源节点—目的节点之间的基础RTT差异会显著大于地面数据中心网络。在这种环境中，RTT差异会严重破坏AIMD拥塞控制算法的带宽公平性。RTT较短的流能够更快地感知可用带宽并完成增长—探测循环，因此在竞争中往往更具侵略性；RTT较大的流由于反馈更慢，速率恢复和拥塞响应都更滞后，最终容易被压制在较低速率。对于星上训练来说，这种不公平不仅意味着平均吞吐量不佳，还可能破坏并行训练所需的阶段同步，使部分长路径的集合通信流长期成为系统瓶颈。

幸运的是，星间拓扑的变化是具有周期性的，这使得基础RTT具有较强的可预测性。在每个时刻，节点位置可由星历信息推算；若路由策略相对稳定，则两节点间路径长度和传播时延可在较高准确度下预测。因此，无须完全依赖被动测量RTT，可以将轨道几何和路由先验直接用于拥塞控制调节。基于这一点，StarRDMA协议在成熟的AIMD拥塞控制框架上加入一个RTT感知的加权因子。由于现有地面数据中心网络、卫星网络的拥塞控制算法较为成熟，因此StarRDMA协议并未提出新的拥塞控制算法，而是对现有算法进行改造。基于卫星拓扑的周期性获得的RTT预测值可以表示为式(3)：

$$T_{\text{prop}} = \sum_{i=1}^{h} \frac{l_i}{c} \tag{3}$$

其中，h为两节点间的转发跳数，l_i为第i跳的链路距离，c为光速。

由每跳的转发处理与排队基线开销t_{proc}，可得预测RTT如式(4)所示：

$$\mathrm{RTT} = 2 \cdot \sum_{i=1}^{h}\left(\frac{l_i}{c} + t_{\mathrm{proc}}\right) \tag{4}$$

在实际系统中，并不要求这一估算绝对精确，只要能够稳定反映不同流之间的相对RTT差异，便足以用于后续的速率加权。补偿RTT差异的基本思路是：让RTT较大的流在加速时更积极，在减速时更克制。本文设RTT_0为一个参考RTT，其对应的增速步长为AI_0，乘性降速因子为β（对应每次降速的速率减少比例）。对某条流f，其预测RTT为$\widehat{\mathrm{RTT}}_f$，则可定义其RTT权重如式(5)所示：

$$\gamma_f = \left(\frac{\widehat{\mathrm{RTT}}_f}{\mathrm{RTT}_0}\right)^{\alpha} \tag{5}$$

其中，$\alpha \in [0,1]$为调节参数，用于控制补偿强度。在此基础上，对AI和MD两个阶段进行加权，分别如式(6)和式(7)所示：

$$\mathrm{AI}{:}\mathrm{AI}_f = \gamma_f \cdot \mathrm{AI}_0 \tag{6}$$

$$\mathrm{MD}{:}\ \beta_f = \beta/\gamma_f \tag{7}$$

对于长RTT流，可通过放大其AI、缩小其MD来减少RTT差异的影响。

2.3.2　语义优先级感知的差异化速率控制

在星间链路带宽有限的背景下，另一个值得关注的问题是：训练系统内部仍然存在大量语义紧迫性不同的流量，各类并行策略产生的集合通信、checkpoint持久化、数据集加载等都共享同一网络，但它们对训练主路径的影响截然不同。若传输层仍采用所有类型的流量公平地共享网络带宽的策略，那么关键路径流量将不可避免地与后台流量争夺稀缺的星间带宽资源，导致训练主路径的数据传输无法被优先保障。一个典型例子是专家并行下的Token分发和聚合。这类流量通常位于计算执行链条的正中央，若Token未送达，则后续计算就无法开始。因此，它的完成时间几乎与GPU的等待时间相等。相比之下，checkpoint流量虽然体量大，但通常不直接阻塞当前微批的前后向计算，因此，即使适度延迟，对训练吞吐量的影响也远小于关键路径通信。这个差异表明，传输层有必要显式引入语义优先级感知的差异化速率控制。为此，StarRDMA协议设计了一种基于优先级的加权增速机制，由应用负责根据语义指定每条流的优先级，并根据优先级为每条流分配差异化的速率。具体而言，StarRDMA协议在速率增长阶段引入优先级权重，例如对于优先级为p的流量来说，其对应的加权权重为w_p，AI增速步长如式(8)所示：

$$\mathrm{AI}(p) = w_p \cdot \mathrm{AI}_0 \tag{8}$$

通过修改增速步长，可使得在相同的增速时间内，高优先级流量占据更多的带宽，满足其高优先级传输的需求。StarRDMA协议采用加权拥塞控制机制，而非多队列调度机制，原因主要是：① 星上交换机的缓存和队列资源更宝贵，复杂多队列调度机制的实现成本较高；② 端侧是流量注入网络的端点，在源端控制流量注入更为自然；③ 训练流量的优先级是动态的，而端侧最了解其语义变化。

2.4　协议硬件实现开销分析

StarRDMA协议作为面向星上训练场景的RDMA类传输协议，其基础数据路径与地面RDMA协议具有共性。其中，QP上下文维护、ACK/NACK处理逻辑、DMA数据搬运以及拥塞速率控制等基础模块，均可沿用地面现有RDMA网卡中的成熟设计。因此，重点分析StarRDMA协议为适配星间链路扰动和训练语义而新增的协议模块带来的资源开销。

StarRDMA协议新增机制的状态开销以及主要计算开销见表2。可以看到，StarRDMA协议引入的额外状态主要来自语义元数据、SACK位图、重传队列以及少量流级

表2　StarRDMA协议新增机制的状态开销以及主要计算开销

模块	新增功能	状态开销	主要计算开销	硬件实现方式	开销判断
语义元数据	在WQE上下文中增加可靠性等级、优先级、截止时间、EC标志等字段	每消息数Byte级状态	$O(1)$字段读取与查表	QP/WQE上下文字段扩展	低
SACK精细反馈	增加记录报文粒度的接收记录，SACK报文的生成以及接收处理	W bit/window，W为乱序容忍窗口内的报文数量	SACK报文接收时逐包扫描携带的报文粒度接收记录，最坏为$O(W)$	SRAM位图	中低
Deadline重传调度	在基础重传队列之上增加服务等级和截止时间调度	$O(L)$，L为待重传报文数	若采用时间桶FIFO，入队/出队近似$O(1)$	多级FIFO以及优先编码器	中
差异化AIMD	在端侧速率控制中加入RTT权重和语义优先级权重	每QP数Byte级权重状态	每反馈事件$O(1)$查表以及一些定点数运算	查表，定点加减/乘移位逻辑	低
选择性RS纠删编码	对时延优先短消息生成冗余报文	编码组元数据$O(k+r)$	编码约$O(krB)$，其中B为报文大小	共享式GF乘加单元	中高

权重信息，整体规模较小，可由片上 SRAM 等常规网卡资源承载。对于计算开销，除 RS 纠删编码外，其余机制在每个报文处理过程中主要涉及字段读取、位图更新、查表、定点加减或队列入出操作，均属于常数级轻量逻辑。相比之下，RS 纠删编码是 StarRDMA 协议中计算开销最高的模块。因此，StarRDMA 协议仅对时延优先短消息选择性启用 EC，而不对集合通信大流默认启用，从而避免在网卡中配置大量全线速 EC 计算单元。

从功耗量级看，当前高性能 RDMA 网卡通常为数十瓦级设备，例如 NVIDIA ConnectX-7 400G 网卡无源铜缆配置下典型功耗约为 25.9 W[26]，Broadcom BCM57608 400G 网卡在 AOC 光缆和 100% Ethernet 流量配置下功耗为 22.3 W[27]。RDMA 网卡的主要功耗来自高速 SerDes、PCIe 总线、DMA 内存访问、片上缓存访问以及协议处理单元，其中高速 SerDes、PCIe 接口以及 DMA 内存访问等 I/O 模块是主要的功耗项。StarRDMA 协议新增模块主要涉及协议处理单元、QP/WQE 上下文的片上 SRAM 访问，不引入新的高速 I/O 通道。结合部分现有卫星平台具备 kW 级乃至更高的太阳能供电能力[12]，上述量级表明高性能 RDMA 网卡在星上供电预算中具有潜在部署空间；StarRDMA 协议新增逻辑的实际资源与功耗开销仍需通过 FPGA/ASIC 原型进一步定量验证。

3 协议仿真与性能分析

本节基于 ns-3 仿真平台对 StarRDMA 协议的关键机制进行仿真评估。评估目标包括：① 验证 StarRDMA 协议的语义感知可靠性机制能否在星间易丢包环境下缩短集合通信完成时间；② 验证 StarRDMA 协议的差异化拥塞控制机制能否缓解 RTT 异构带来的性能劣化，以及优先保障训练关键路径流量的效果。

3.1 仿真环境说明

3.1.1 仿真实现

本文基于 ConWeave ns-3 仿真框架[28]实现 StarRDMA 协议的传输逻辑。ConWeave 已提供了基础的 RDMA 传输逻辑，本文在该仿真框架上实现了 StarRDMA 协议的关键机制。需要说明的是，本文在仿真中并未实现完整的字节级 RS 编解码器，而是采用仿真层面的 RS(k+r, k)可恢复性抽象。具体而言，发送端按照每 k 个数据包附加 r 个冗余包的方式建模纠删编码带来的额外传输开销，当某个编码组内已有不少于 k 个包成功到达接收端时，接收端即认为该组可由 RS 解码恢复。在下述仿真中，本文配置的 RS 编码参数为 RS(10, 8)，即 k=8, r=2。

除协议机制外，本文还构建了一个动态周期运行的卫星紧凑编队拓扑，用于刻画星间相对运动引起的传播时延变化和链路质量波动。该拓扑受 Google “Project Suncatcher” 计划[12]中紧凑编队互连设计的启发，但由于该计划并未公开其紧凑编队拓扑的完整几何模型及轨道参数，因此本文采用一个近似模型模拟该拓扑的卫星运动规律。具体而言，卫星节点在逻辑邻接关系上构成 2D-Mesh，但其几何布局并非规则矩形。为模拟紧凑编队中卫星相对位置的周期变化，将 2D-Mesh 卫星节点初始嵌入椭圆形编队平面中，并基于 Hill-Clohessy-Wiltshire（HCW）相对运动模型构造其周期动态。设卫星 i 和 j 在初始编队坐标系下的相对位置如式(9)所示：

$$\Delta x_0 = x_i - x_j,\ \Delta y_0 = y_i - y_j \tag{9}$$

在仿真时间 t 下，周期相位定义如式(10)所示：

$$\theta(t) = 2\pi \cdot \frac{t \bmod T_{\text{orbit}}}{T_{\text{orbit}}} \tag{10}$$

二者的瞬时相对位置更新如式(11)所示：

$$\begin{aligned} \Delta x(t) &= \Delta x_0 \cos\theta(t) + 0.5\Delta y_0 \sin\theta(t) \\ \Delta y(t) &= -2\Delta x_0 \sin\theta(t) + \Delta y_0 \cos\theta(t) \end{aligned} \tag{11}$$

对应的瞬时星间距离如式(12)所示：

$$d_{ij}(t) = \sqrt{\Delta x(t)^2 + \Delta y(t)^2} \tag{12}$$

仿真过程中，链路邻接关系固定，但每条星间链路的传播时延根据瞬时距离动态更新。

默认 2D-Mesh 规模为 24×24，共包含 576 个卫星节点，每个卫星节点被抽象为一个星上计算超节点，其内部可通过高带宽机内互连（如 NVLink）连接多个 GPU。由于本文关注的是星间 RDMA 协议的传输层行为，因此不进一步建模超节点内部的 GPU 拓扑。当每个卫星节点包含 72 个 GPU 时，该配置对应 4.15 万张 GPU 的星上分布式训练规模。本文将直接相邻卫星节点在周期内的平均距离校准为 3 km，以近似紧凑编队飞行下的星间邻接关系。所有星间链路均配置为 100 Gbit/s 全双工链路，并采用 dimension-order 路由进行确定性转发。

本文未采用 Walker 星座作为评估拓扑，原因在于 Walker 星座的设计目标与 StarRDMA 协议的目标场景并不一致。Walker 星座通常面向全球覆盖型卫星网络，卫星分布在多个轨道面上，星间路径可能跨越较长地理距离，因此更适合评估面向广域互联网服务的传输协议。相比之下，StarRDMA 协议面向紧凑卫星编队内部的大规模分

布式训练场景，卫星被有意布置在较短星间距离内，以提供高带宽、低时延的计算节点互连。因此，本文采用紧凑编队下的2D-Mesh近似模型，重点刻画与StarRDMA协议传输机制最相关的动态因素，包括星间距离变化、传播时延波动以及距离相关的链路丢包。

3.1.2　对比基线方案

本文在ns-3中实现了StarRDMA协议的语义感知可靠性机制和差异化拥塞控制机制，在下面的仿真实验中将对这两个机制进行仿真评估。为了分析不同机制的影响，分别选取如下基线方案。① 可靠性机制评估基线：采用直接迁移至星间网络的RoCEv2 Go-back-N丢包恢复机制。② 拥塞控制机制评估基线：采用经过参数调优的数据中心量化拥塞通知（data center quantized congestion notification, DCQCN）算法[29]。对于可靠性机制评估，StarRDMA协议与基线统一采用同一组DCQCN参数，以避免拥塞控制差异干扰恢复机制本身的比较；对于拥塞控制机制评估，则统一启用StarRDMA协议的基础丢包恢复机制，且所有流量均采用严格可靠服务，以说明拥塞控制差异带来的影响。

3.1.3　评估工作负载

本文主要评估两类训练中最典型的集合通信模式：AlltoAll和AllReduce。其中，AllReduce的实现采用适配2D-Mesh拓扑的分阶段算法（参考文献[30]），以代表大规模并行训练中的梯度同步；AlltoAll则用于刻画专家并行、张量重排等需要全互连交换的数据搬运场景。

在可靠性机制实验中，本文在链路层注入非拥塞丢包，以评估不同链路质量条件下StarRDMA协议的恢复能力。具体而言，本文考虑两类丢包模型：① 独立随机丢包模型，用于模拟由链路误码引起的随机包损坏或丢弃，基础丢包率设置为0.1%、1.0%、5.0%和10.0%。在基础丢包率的基础上，本文进一步将丢包概率与链路距离相关联：当星间相对运动使链路距离增大时，链路预算余量下降，丢包概率相应增大；当链路距离缩短时，丢包概率降低。该模型用于刻画星间相对运动对链路质量的影响。② Gilbert-Elliott突发丢包模型，用于刻画链路短时不稳定、光链路指向扰动或信道质量下降导致的短时连续丢包现象。在该模型中，链路在Good和Bad两个状态之间转移，Good状态下丢包率较低，Bad状态下丢包率较高，从而形成具有时间相关性的突发丢包。在动态链路下，本文同样将链路距离引入状态转移过程，使较长链路更容易进入Bad状态或维持较高的突发丢包概率。通过同时引入距离相关随机丢包和距离调制突发丢包，本文能够评估StarRDMA协议在独立误码、连续丢包以及星间距离变化共同作用下的恢复性能。除特别说明外，本文图中标注的丢包率表示链路在一个轨道周期内的目标平均丢包率。

在拥塞控制实验中，本文构造两类场景评估StarRDMA协议的差异化速率控制能力。一是不同跳数的竞争流场景，即同时注入路径长度不同、但在部分链路上发生竞争的RDMA流。由于不同跳数对应不同的传播时延和反馈时延，该实验用于评估RTT异构条件下的带宽竞争行为，并验证RTT加权增速机制是否能够缓解长RTT流速率恢复较慢的问题。二是后台大流与周期性突发AlltoAll共存的混合负载场景，用于模拟星上训练过程中低优先级数据搬运与关键路径集合通信并发的情况。其中，后台大流持续占用链路带宽，周期性AlltoAll代表对完成时间敏感的前景通信。该实验用于评估语义优先级加权增速机制是否能够在拥塞恢复过程中优先保护关键路径流量，从而降低前景AlltoAll的平均完成时间和尾部时延。

3.2　可靠性机制评估

3.2.1　基础丢包恢复机制的有效性

本文选择AlltoAll和AllReduce作为负载，分别比较StarRDMA协议与RoCEv2 Go-back-N机制在不同丢包率下的表现。为避免可靠性分级因素干扰，本实验中所有消息均采用严格可靠服务，并且StarRDMA协议不开启纠删编码。同时，两种方案统一采用经过调优的DCQCN拥塞控制参数，因此性能差异主要来自丢包检测与恢复粒度的不同。实验采用第3.1.3节中定义的距离相关随机丢包模型和距离调制Gilbert-Elliott突发丢包模型；除特别说明外，横轴标注的丢包率表示一个轨道周期内的目标平均丢包率。

图5给出了独立随机丢包模型下AlltoAll与AllReduce在不同丢包率下的平均完成时间和第99百分位（99th percentile, P99）完成时间。可以看到，在低丢包率（如0.1%）时，两种方案的差距相对有限，随着丢包率升高，Go-back-N的完成时间迅速恶化，而StarRDMA协议的完成时间增长趋势明显更缓。具体而言，在5.0%的丢包率下，StarRDMA协议可将AlltoAll的平均完成时间和P99完成时间分别降低52.08%和56.25%；在10.0%的丢包率下，这一优势进一步扩大到60.61%和64.22%。在AllReduce上也观察到相同趋势：在5.0%的丢包率下，StarRDMA协议可分别降低平均完成时间和P99完成时间

48.33%和50.93%；在10.0%的丢包率下，对应降幅进一步提升至59.72%和61.61%。

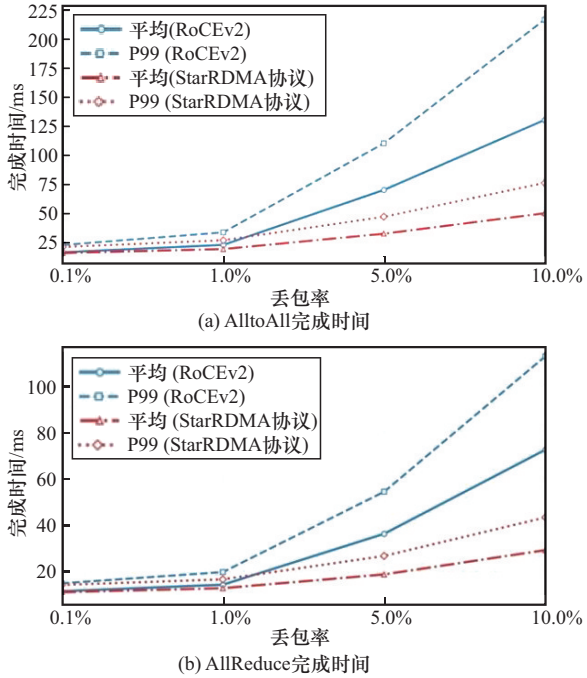


图5　独立随机丢包模型下不同丢包率对应的集合通信完成时间

图6给出了Gilbert-Elliott突发丢包模型下不同丢包率对应的集合通信完成时间。相比于独立随机丢包模型，突发丢包模型对Go-back-N更不友好：连续丢包更容易触发乱序、回退和重复发送，使恢复开销和尾部等待时间被进一步放大。因此，在该模型下StarRDMA协议的优势更加明显。以丢包率为5.0%为例，StarRDMA协议可将AlltoAll的平均完成时间和P99完成时间分别降低63.14%和67.44%，将AllReduce的平均完成时间和P99完成时间分别降低62.10%和63.54%；在10.0%的丢包率下，上述降幅分别进一步提升至AlltoAll的71.02%和72.83%，以及AllReduce的69.57%和71.06%。

造成上述差异的根本原因在于，两种恢复机制对丢包的处理粒度不同。Go-back-N在发生丢包后，往往需要回退并重复发送一段连续报文，在高RTT环境下这会放大恢复等待时间，并额外占用本就稀缺的星间带宽；在突发丢包场景下，这种回退式恢复还会被连续丢包进一步放大。相比之下，StarRDMA协议通过SACK反馈精确定位缺失报文，仅重传真正丢失的数据分组，从而显著降低了误码丢包和突发丢包带来的级联恢复开销。对于同步型集合通信，这种恢复粒度的差异会直接体现在完成时间上，尤其是尾部完成时间。

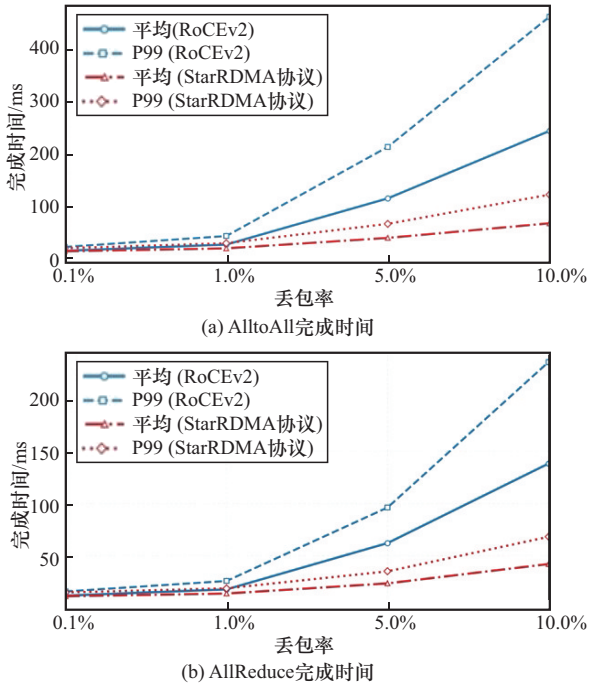


图6　Gilbert-Elliott突发丢包模型下不同丢包率对应的集合通信完成时间

3.2.2　可靠性分级服务的必要性

第3.2.1节说明了基于SACK的精细恢复优于Go-back-N。本节进一步评估在采用相同基础恢复机制的情况下，若仍然对所有训练流量统一采用严格可靠服务，是否会引入额外恢复负担，以及StarRDMA协议的分级可靠性设计能否进一步改善集合通信完成时间。选择AlltoAll作为代表性集合通信负载，并比较以下两种方案：统一可靠服务——所有消息均采用严格可靠服务；分级可靠服务——根据消息语义采用严格可靠、部分可靠和时延优先三类可靠性服务。默认情况下，三类流量的消息数比例设置为严格可靠：部分可靠：时延优先 = 60：30：10。其中，严格可靠流量要求完整交付；部分可靠流量允许在超过截止时间后跳过部分恢复；时延优先流量主要对应对完成时间敏感的小消息，并通过冗余编码降低重传等待开销。除可靠性服务等级外，其余网络与拥塞控制设置与前述实验保持一致，统一采用DCQCN算法。

图7给出了不同丢包率下的集合通信完成时间。结果表明，在低丢包率条件下，分级可靠性已经能够降低恢复开销；随着丢包率升高，分级可靠方案相比于完全可靠方案展现出更明显的性能优势。对于严格可靠流量，在10.0%的丢包率下，分级可靠方案可将平均完成时间和P99完成时间分别降低54.41%和54.35%。对于时延优先

流量，由于其消息规模更小且可利用冗余编码减少丢包后的恢复等待，因此，在10.0%的丢包率下，分级可靠方案可将平均完成时间和P99完成时间分别降低75.56%和68.89%。

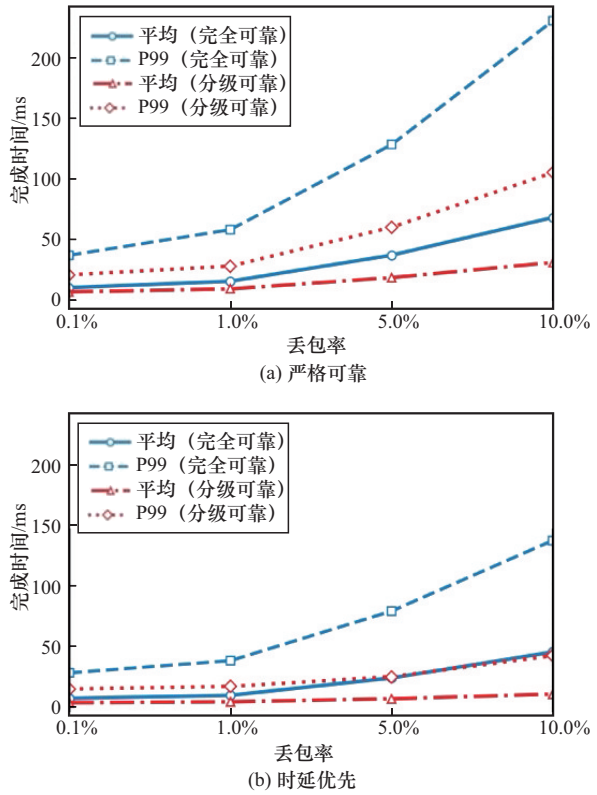


图7　不同丢包率下的集合通信完成时间

进一步地，图8给出了1.0%丢包率下部分可靠服务的截止时间敏感性结果。图中横轴表示部分可靠服务截止时间相对于理论完成时间的放宽倍数。可以看到，当截止时间放宽倍数从2倍增加到8倍时，超时跳过报文数逐渐减少，说明协议会尝试恢复更多数据块；但与此同时，部分流需要等待更长的恢复过程，导致平均完成时间和P99完成时间上升。这一结果反映了部分可靠服务的核心权衡：更激进的截止时间可以减少恢复等待、缩短集合通信完成时间，但会带来更多受控跳过；更宽松的截止时间可降低跳过代价，但会使少数慢流拖慢整体完成时间。

仿真结果表明，对星上训练流量统一采用最保守的严格可靠语义并非最优。其根本原因在于，不同消息在训练过程中的价值并不相同：部分消息必须完整恢复，而另一些消息则允许受控丢失或更适合通过冗余编码避免长时间重传等待。若按统一严格可靠语义处理，协议会将恢复带宽和恢复等待时间平均施加到所有消息上，从而导致真正位于关键路径上的流量无法及时完成。StarRDMA协议通过将训练语义映射到不同可靠性服务等级，在保证关键数据完整性的同时，减少非关键恢复开销，从而降低集合通信完成时间，尤其是可降低高丢包率场景下的尾部完成时间。

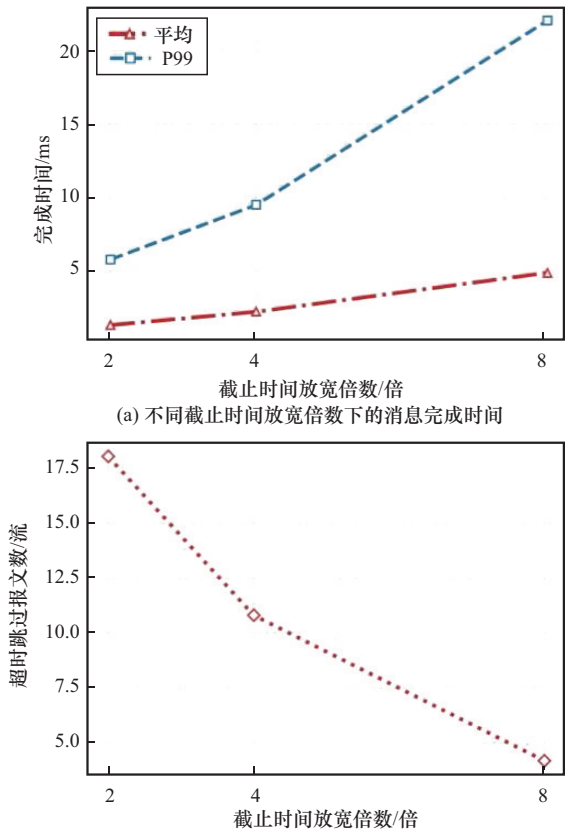


图8　1.0%丢包率下部分可靠服务的截止时间敏感性结果

3.3　拥塞控制机制评估

接下来评估StarRDMA协议的两项拥塞控制机制：基于轨道RTT预测的加权AIMD，以及基于语义优先级的差异化增速。除特别说明外，本节所有流量均采用严格可靠服务，并统一启用StarRDMA协议的基础丢包恢复机制以隔离拥塞控制本身的作用。

3.3.1　基于RTT预测的加权AIMD

本文通过构造一个多打一的Incast场景，使多条流竞争同一接收端或共享瓶颈链路，并通过改变源节点与目的节点之间的跳数来控制路径RTT。不同跳数下长RTT流的完成时间如图9所示，其中，横轴为流的跳数，纵轴为完成时间，对比方案为不引入RTT补偿的DCQCN算法。

图 9 表明，在基线 DCQCN 算法下，随着跳数增加，流完成时间近似单调上升；在 StarRDMA 协议下，完成时间随跳数增加的恶化趋势明显减弱。特别是在中长路径（如 8 到 24 跳）区间，StarRDMA 协议相比 DCQCN 算法将 P99 完成时间最高降低 42.86%。该结果说明，RTT 异构确实会在星间多跳网络中放大传统 AIMD 类拥塞控制的不利影响。对于基础 RTT 更短的流，由于反馈回路更快，其速率恢复和带宽探测更积极；长路径流由于反馈滞后，往往难以及时占据瓶颈带宽，从而在竞争中持续落后。StarRDMA 协议利用轨道几何与路由信息提供的 RTT 先验信息，对增减速过程进行加权，使长 RTT 流在加速阶段更积极、在减速阶段更克制，从而缩小了不同跳类流之间的性能差距。

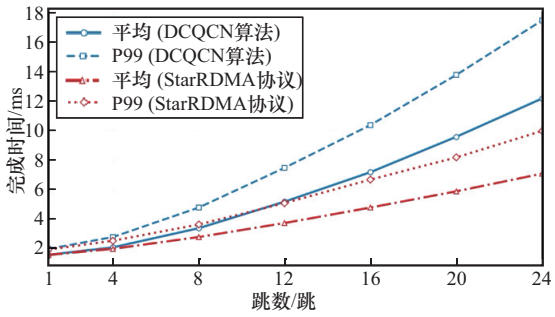


图 9　不同跳数下长 RTT 流的完成时间

3.3.2　语义优先级感知的差异化增速

本文构造两类共存流量：一类是后台流量，即持续存在的 checkpoint 类型的长流，用于模拟训练中不直接阻塞当前微批执行的后台带宽消耗；另一类是前景流量，即周期性突发的 AlltoAll，用于模拟训练关键路径上的同步通信。对比方案为采用统一增速策略的 DCQCN 算法。StarRDMA 协议为前景 AlltoAll 赋予更高优先级，在增速阶段使用更大的加权步长。

图 10 给出了不同后台流量负载下前景 AlltoAll 流量的完成时间。结果显示，当后台长流持续占用链路带宽时，采用统一增速策略的 DCQCN 算法会逐渐拉长前景 AlltoAll 的完成时间；相比之下，StarRDMA 协议能够为前景关键路径流量提供更积极的增速，使其平均完成时间和尾部完成时间保持在更低水平。随着后台流量负载从 20% 增加到 60%，StarRDMA 协议的收益逐渐扩大。在 60% 的后台流量负载下，StarRDMA 协议可将前景 AlltoAll 的平均完成时间降低 32.20%，P99 完成时间降低 35.71%。这说明 StarRDMA 协议通过在端侧拥塞控制过程中直接引入语义优先级，使关键路径流量在带宽探测时更快获得发送机会，从而降低后台流量对同步通信完成时间的影响。

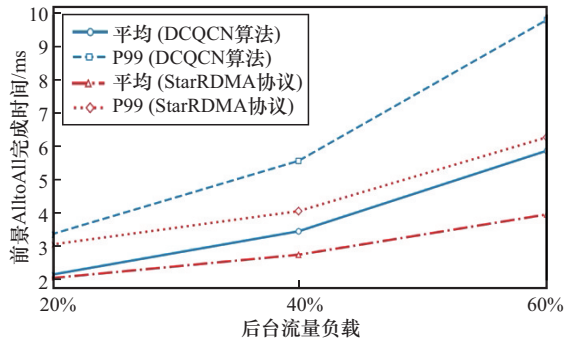


图 10　不同后台流量负载下前景 AlltoAll 流量的完成时间

4　结束语

本文研究了星上大模型分布式训练中的高性能传输问题，基于传统卫星传输协议难以满足训练通信对高吞吐量、低尾时延和稳定完成时间的要求，以及地面 RDMA 协议所依赖的网络假设在星间环境中难以成立的现状，提出一种面向星上训练的语义感知 RDMA 传输协议——StarRDMA 协议，其通过多等级可靠性、选择性纠删编码、截止时间感知重传，结合轨道 RTT 预测和语义优先级的差异化拥塞控制机制，可在星间时变网络中更好地支持训练关键路径通信。总体而言，将训练语义显式引入传输层设计，是构建星上大模型训练高性能互联的重要方向。未来工作中，将进一步构建 StarRDMA 协议的现场可编程门阵列（FPGA）原型，对协议新增机制的逻辑资源占用、片上存储开销及功耗进行定量评估。

参考文献

[1] Kaplan J, McCandlish S, Henighan T, et al. Scaling laws for neural language models[PP]. arXiv preprint, 2020, arXiv: 2001.08361.

[2] Yang A, Li A F, Yang B S, et al. Qwen3 technical report[PP]. arXiv preprint, 2025, arXiv: 2505.09388.

[3] DeepSeek AI. DeepSeek-V4: towards highly efficient million-token context intelligence[R]. 2026.

[4] Jiang Z, Lin H, Zhong Y, et al. MegaScale: scaling large language model training to more than 10,000 GPUs[C]//21st USENIX Symposium on Networked Systems Design and Implementation (NSDI 24). 2024: 745-760.

[5] Gangidi A, Miao R, Zheng S B, et al. RDMA over Ethernet for dis-

tributed training at meta scale[C]//Proceedings of the ACM SIGCOMM 2024 Conference. New York: ACM, 2024: 57-70.

[6] 张慧峰, 刘宁春, 龙卫平, 等. 面向国产算力的超大规模智算集群网络: 关键挑战、技术途径与发展趋势[J]. 电信科学, 2025, 41(12): 1-26.
Zhang H F, Liu N C, Long W P, et al. Hyperscale intelligent computing cluster networks for domestic computing power: critical challenges, technical pathways, and future trends[J]. Telecommunications Science, 2025, 41(12): 1-26.

[7] 林子涵. 拆解 AI 背后的“国际能耗账单”(环球热点)[N]. 人民日报海外版, 2026-02-26.
Lin Z H. Breaking down the “international energy bill” behind AI (global hotspot)[N]. People’s Daily Overseas Edition, 2026-02-26.

[8] 胡媛. 算力电力消耗究竟有多大？[N]. 通信产业报,2024-10-28.
Hu Y. How much electricity does computing power really consume?[N]. Communications Industry News, 2024-10-28.

[9] IEA. Energy and AI: energy demand from AI[R]. 2025.

[10] 张哲宇, 李泓辛, 蓝天翼, 等. 太空数据中心的需求分析与国外发展现状[J]. 电讯技术, 2026, 66(5): 865-880.
Zhang Z Y, Li H X, Lan T Y, et al. Space data centers: requirement analysis and overseas development status[J]. Telecommunication Engineering, 2026, 66(5): 865-880.

[11] Bargatin I, Jin D G, Alansari Z, et al. Tether-based architecture for solar-powered orbital AI data centers[PP]. arXiv preprint, 2025, arXiv: 2512.09044.

[12] Arcas B A Y, Beals T, Biggs M, et al. Towards a future space-based, highly scalable AI infrastructure system design[PP]. arXiv preprint, 2026, arXiv: 2511.19468.

[13] 刘波, 杨孟飞, 李明, 等. 太空计算中心构建及运营技术研究[J]. 宇航学报, 2025, 46(12): 2462-2471.
Liu B, Yang M F, Li M, et al. Technologies for construction and operational management of space computing centers[J]. Journal of Astronautics, 2025, 46(12): 2462-2471.

[14] 张超, 时千舒, 孟斌斌, 等. 超越创新悖论: 低成本推动商业航天高质量发展的实践与思考[J]. 中国科学院院刊, 2025, 40(11): 1902-1913.
Zhang C, Shi Q S, Meng B B, et al. Beyond innovation paradox: Practices and reflections on promoting high-quality development of commercial aerospace with low costs[J]. Bulletin of Chinese Academy of Sciences, 2025, 40(11): 1902-1913.

[15] Kang S Y, Jo M S, Choi J Y, et al. Cost effectiveness of reusable launch vehicles depending on the payload capacity[J]. Aerospace, 2025, 12(5): 364.

[16] 王浩楠, 刘峻峰, 南卓江, 等. 星间激光通信关键技术与展望[J]. 激光与光电子学进展, 2025, 62(9): 0900006.
Wang H N, Liu J F, Nan Z J, et al. Key technology and prospect of inter-satellite laser communication[J]. Laser & Optoelectronics Progress, 2025, 62(9): 0900006.

[17] 陈山枝, 范志文, 金家德, 等. 卫星互联网星间激光通信的分析及建议[J]. 电信科学, 2024, 40(2): 1-10.
Chen S Z, Fan Z W, Jin J D, et al. Analysis and suggestions on inter-satellite laser communication of satellite Internet[J]. Telecommunications Science, 2024, 40(2): 1-10.

[18] 刘鹏英, 何颖, 田路云, 等. 星载人工智能芯片应用现状与产业发展研究[J]. 飞控与探测,1-11.
Liu P Y, He Y, Tian L Y, et al. Current status of applications and industrial development of spaceborne artificial intelligence chips [J]. Flight Control & Detection, 1-11.

[19] 管晶晶. 全球首次！通用大模型太空在轨部署成功 [N]. 科技日报, 2026-01-27.
Guan J J. A world first! general-purpose large model successfully deployed in orbit [N]. Science and Technology Daily, 2026-01-27.

[20] Lee A. How starcloud is bringing data centers to outer space[EB]. 2025.

[21] 新华社. 我国整轨互联太空计算星座“三体计算星座”正式进入组网阶段[N]. 新华网, 2025-05-14.
Xinhua News Agency. China’s in-orbit interconnected space computing constellation, the “three-body computing constellation” officially enters the networking phase [N]. XinhuaNet, 2025-05-14.

[22] Pultarova T. SpaceX plan for 1 million orbiting AI data centers could ruin astronomy, scientists say[EB]. 2026.

[23] InfiniBand Trade Association. InfiniBand architecture specification: volume 1, release 1.2.1[S]. 2007.

[24] Guo C X, Wu H T, Deng Z, et al. RDMA over commodity Ethernet at scale[C]//Proceedings of the 2016 ACM SIGCOMM Conference. New York: ACM, 2016: 202-215.

[25] Chen X Q, Vargaftik S, Ben Basat R. When ML training cuts through congestion: just-in-time gradient compression via packet trimming[C]//Proceedings of the 23rd ACM Workshop on Hot Topics in Networks. New York: ACM, 2024: 177-185.

[26] NVIDIA. NVIDIA ConnectX-7 adapter card specifications[EB]. 2026.

[27] Broadcom. BCM957608 400G Ethernet adapters data sheet: 957608-DS100[EB]. 2026.

[28] Conweave-project. conweave-ns3[CP]. 2023.

[29] Zhu Y B, Eran H, Firestone D, et al. Congestion control for large-scale RDMA deployments[C]//Proceedings of the 2015 ACM Con-

ference on Special Interest Group on Data Communication. New York: ACM, 2015: 523-536.

[30] Kumar S, Jouppi N. Highly available data parallel ML training on mesh networks[PP]. arXiv preprint, 2020, arXiv: 2011.03605.

[作者简介]



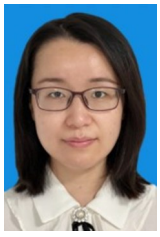
黄俊凯（2000– ），男，北京邮电大学硕士生，主要研究方向为卫星通信网络、智算中心网络。



潘恬（1987– ），男，博士，北京邮电大学教授、博士生导师，主要研究方向为云数据中心网络、低轨卫星网络、高速可编程网络设备等。



杨祎凌（2001– ），男，北京邮电大学硕士生，主要研究方向为云网络网关。



张娇（1986– ），女，博士，北京邮电大学教授、博士生导师，主要研究方向为智算中心网络、数据中心网络、网络传输控制协议等。



黄韬（1980– ），男，博士，北京邮电大学教授、博士生导师，主要研究方向为未来网络体系结构、软件定义网络、确定性网络等。