



专题：智算光互联

## 面向智算中心间互联的光网络关键技术研究

李韞瑄<sup>1,2</sup>, 杨亚萍<sup>1,2</sup>, 涂佳一<sup>1,2</sup>, 顾仁涛<sup>1,2</sup>, 纪越峰<sup>1,2</sup>

(1. 北京邮电大学信息与通信工程学院, 北京 100876;

2. 北京邮电大学信息光子学与光通信全国重点实验室, 北京 100876)

**摘要:** 智算中心作为大模型应用的关键算力基础设施, 其高效运行依赖于高性能的光网络传输底座。然而, 智算中心间光网络面临包括高实时性、高突发性和高可靠性的三重互联需求带来的诸多技术挑战。基于此, 需要以高实时资源分配技术, 助力智算光网络资源分配突破局部调度限制, 以应对业务传输与调度时延挑战; 以自适应协同优化技术, 引导网络从被动调整向主动协同持续发展, 以应对业务动态高突发性挑战; 以前摄性故障恢复技术, 推动网络从被动恢复故障向主动干预有序演进, 以应对智算光网络的高可靠性挑战。展望未来, 高实时大规模调度技术、算网深度协同技术和智算数字孪生技术, 将为智算中心间互联的进一步发展提供新的机遇。

**关键词:** 智算中心; 资源分配; 网络优化; 故障恢复

**中图分类号:** TP393

**文献标志码:** A

**doi:** 10.11959/j.issn.1000-0801.2025106

## Research on key technologies of optical networks for interconnection between artificial intelligent data centers

LI Yunxuan<sup>1,2</sup>, YANG Yaping<sup>1,2</sup>, TU Jiayi<sup>1,2</sup>, GU Rentao<sup>1,2</sup>, JI Yuefeng<sup>1,2</sup>

1. School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

2. State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

**Abstract:** As a critical computing infrastructure for large-scale model applications, the efficient operation of artificial intelligence data center relies on high-performance optical network transmission infrastructure. However, the optical networks interconnecting artificial intelligence data center face numerous challenges in meeting the demands of high real-time, high burstiness, and high reliability. Based on this, real-time resource allocation aims to overcome localized scheduling limitations in optical networks interconnecting artificial intelligence data center and reduce transmission

收稿日期: 2025-02-10; 修回日期: 2025-04-08

通信作者: 顾仁涛, rentaogu@bupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.62271078); 北京市自然科学基金资助项目 (No.L247024)

**Foundation Items:** The National Natural Science Foundation of China (No. 62271078), Beijing Natural Science Foundation (No. L247024)



and scheduling delays. Adaptive and collaborative optimization enables a continuous progression from passive adjustment to active collaboration in response to dynamic traffic bursts. Proactive failure recovery aims to achieve an orderly evolution from passive restoration to active intervention for reliability in optical networks interconnecting artificial intelligence data center. Future developments in large-scale real-time scheduling, deep computing-network convergence, and AI digital twins will drive new advancements in artificial intelligence data center interconnections.

**Key words:** artificial intelligent data center, resource allocation, network optimization, failure recovery

## 0 引言

基于大语言模型的生成式人工智能技术，正引领科技革命与产业变革进入新阶段，各行各业从“+AI”到“AI+”的转变已势不可挡<sup>[1]</sup>。随着如GPT系列、LLaMA系列、DeepSeek系列等大规模模型参数的不断增多，AI大模型在语言理解、逻辑推理和复杂问题分析等方面得到了显著提升。然而，训练这些超大规模模型给智能计算基础设施带来了前所未有的挑战。与传统的数据中心和超算中心相比，面向图形处理器（graphics processing unit, GPU）互联的智算中心可以更为高效地满足日益增长的算力需求。《中国算力发展报告（2024年）》显示，截至2023年年底，全球智算市场同比增长130%以上，中国智算市场同比增长超过60%。

2022年1月，国务院印发《“十四五”数字经济发展规划》，提出加快建设信息网络基础设施，有序推进基础设施智能升级，加快实施“东数西算”工程的要求。随着“东数西算”工程的推进，数据和任务不再局限于智算中心的单点作战，而会在跨地域的智算中心间流转，持续聚焦于跨地域的算力协同。智算中心的跨地域协同主要由以下两个要素驱动。

- 由小规模到大规模。我国单点智算中心规模普遍偏小，规模为100~300 PFLOPS的小型智算中心占比超70%，而规模超过1 EFLOPS的大型智算中心仅占25%<sup>[2]</sup>。智算中心间互联可以对小规模智能计算基础设施进行整合，不需要构

建大规模集约型智算中心，就能够训练更大的模型。

- 由碎片化到集中式。尽管智算中心已经投入大规模应用，然而面向非专用任务的智算中心，仍然存在算力闲置现象，传统模式下的智算中心GPU利用率甚至低于30%<sup>[3]</sup>。《2021—2022全球算力指数评估报告》显示，智算中心的理想利用率至少为80%才能达成显著经济效益。智算中心间互联可以对碎片化算力进行整合，从而提升智算中心的利用率，实现经济效益的提升。

因此，推动智算中心间互联成为提升算力和经济效益的重要方式。在大规模训练场景下，智算中心内部的机外集合通信将产生大量通信需求。例如，千亿级参数的大模型并行训练所产生的集合通信数据将达到数百GB量级。面向智算中心间互联场景，汇聚后的集合通信数据量将进一步增长。同时，网络拥塞和丢包将会严重影响GPU计算效率，因此，智算中心间互联要求传输网络具备高带宽、无损传输能力。光网络由于其大带宽和无损传输特性，成为承载智算中心传输的网络底座。同时，IP层的灵活性和动态调整能力可以在IP/光网络下提供流量疏导，跨层资源分配有助于提升网络利用率、减少网络拥塞。因此，IP/光网络成为支持智算业务需求高带宽、无损传输需求的必要承载场景，共同构建起智算光网络。智算中心间互联网络结构如图1所示。

本文首先分析了智算中心间互联对现有光网

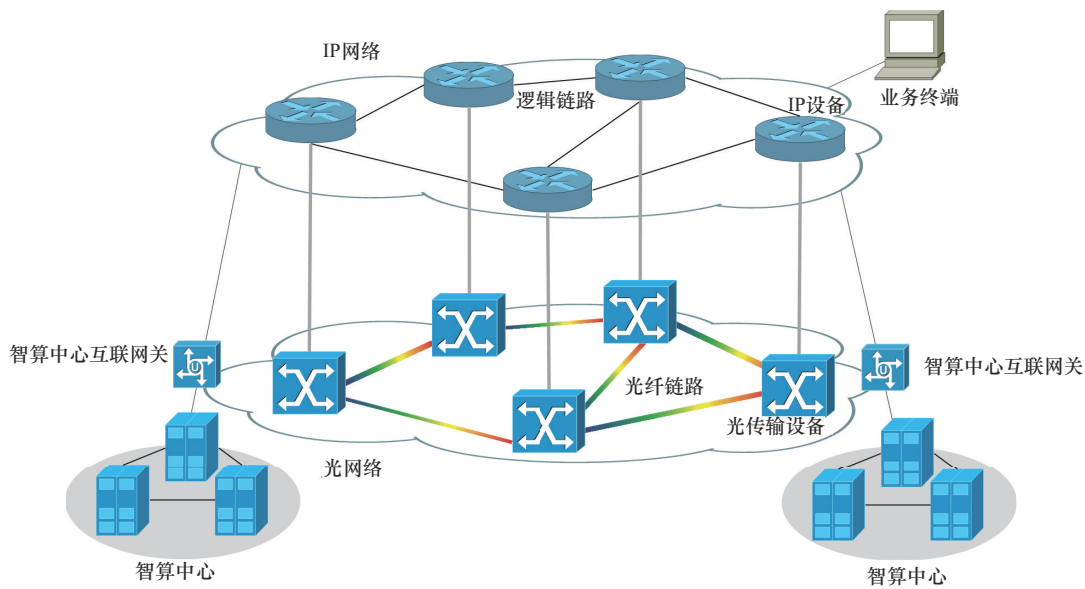


图1 智算中心间互连网络结构

络带来的3个新挑战，然后从智算光网络资源分配技术、智算光网络自适应优化技术以及智算光网络前摄性故障恢复技术3个方面介绍相关研究工作，最后讨论了智算中心间互联光网络的未来研究方向。

## 1 智算中心间互联的光网络关键技术挑战

智算中心的发展对光网络提出了高实时性、高突发性和高可靠性三大挑战。针对这些挑战，已有技术在光网络资源分配、网络优化与故障恢复等方面展开创新，以满足智算中心的业务需求。

### 1.1 高实时性挑战—全局规划的资源分配技术

智算中心的高实时训练过程使智算业务具有高实时性传输需求，这给光网络的资源分配技术带来了新的挑战，特别是在线学习与增量训练等特定场景对高实时性提出了更高的要求。例如，针对自动驾驶和工业互联网等场景<sup>[4]</sup>，模型需要实时适应环境变化，通过在线学习与增量训练快速更新模型参数；针对跨智算中心协同训练场景，多个智算中心需要实时同步数据，以确保模型的一致性和训练效率。智算中心间的长距离传

输会提高传输时延，而高实时性要求的光网络从本质上需要同时满足传输的实时性和调度的实时性。因此，如何突破局部调度限制，以高实时光网络传输和调度技术为基础，进一步利用跨层组网（如IP与光混合组网）全局规划的传送能力优势，充分发挥网络传送潜力，成为应对智算业务高实时性挑战的研究重点。

### 1.2 高突发性挑战——主动协同的网络优化技术

智算业务的周期性传输特性导致其流量具有高突发性，这给光网络的网络优化技术带来了新的挑战。AI大模型的训练过程通常需要对数据进行周期性处理与迭代，使得集合通信数据的产生存在一定的周期性特征，加之智算节点间单次数据传送量巨大，导致其流量突发性明显，易于出现传输的流量总和短时超出网络承载能力的情况，从而引发拥塞。为了应对突发流量，网络中通常会预留大量资源，这使得网络中存在碎片化资源。因此，进一步推动从被动调整迈向主动协同，利用主动的“网”与“端”的协同优化来提升网络性能，成为应对智算业务高突发性挑战的研究重点。



### 1.3 高可靠性挑战——主动干预的故障恢复技术

智算业务的无损传输需求要求光网络具备高可靠性，这给光网络的故障恢复技术提出了新的挑战。面向智算中心间传输场景，为满足高性能算力互联需求，承载网必须具备长距离无损确定性传输能力。尽管光网络在正常运行时可以视作无损传输网络，但是IP层或光层设备故障仍然会导致光网络上业务的传输受到影响，从而带来算力损失。远程直接存储器访问（remote direct memory access, RDMA）因其无损传输的需求<sup>[5]</sup>，对网络性能有着极高的要求。实验统计表明，仅0.1%的网络丢包率就会带来50%的算力损失，因此，针对智算中心间互联的光网络，故障恢复技术的研究显得至关重要。如何能够利用对故障的主动干预，打破传统的“故障发生—手动排查—故障恢复”的被动恢复模式，成为实现光网络高可靠性的研究重点。

## 2 智算光网络资源分配技术

智算业务的高实时需求使得网络需要同时满足高实时传输需求以及高实时调度需求。在跨智算中心协同训练过程中，多个智算中心需要在短时间内传输大量数据，大带宽业务的实时传输对光网络资源分配提出了挑战。基于此，研究人员通过光网络的单链路扩容和多链路协同传输，共同实现大带宽业务在光网络的快速传输，以响应智算业务的实时传输需求。同时，考虑IP/光网络的多层网络资源，通过对多层网络进行全局考虑，可以将多个低速率流量疏导到同一光通道，从而优化网络资源利用。随着网络规模的扩大，高实时调度难度进一步提升，大规模网络下的快速并行资源分配技术成为解决高实时挑战的关键技术之一。

### 2.1 高实时传输技术

为实现大带宽智算业务的高实时传输，光网络需要为智算业务提供大带宽支持。在单链路

上，多波段扩容技术的出现为高实时传输提供了技术支撑。为打破单链路瓶颈，多路径协同传输为高实时传输提供了新的思路。同时，IP/光网络的跨层流量疏导技术为智算业务的实时传输提供了全局规划的资源协调基础。此外，当智算业务存在确定性传输时延需求时，为保障智算中心间业务的传输，在进行资源分配时需要路径长度进行约束。

#### (1) 多波段光层资源分配技术

在尽可能充分利用现有光纤部署的前提下，增加标准C波段以外的波段（如O、E、S和L波段）成为提高光网络容量的优选技术路线，也成为适配智算业务高实时传输需求的关键技术。

然而，在引入多波段高维度的信道场景后，光信号的受激拉曼散射等宽带非线性效应显著增强，传输质量对光网络资源调控的影响大幅加剧<sup>[6]</sup>，光纤物理特性、网络业务分布、光信号调制方式以及频谱资源分配的耦合性加强，传统的波分复用（wavelength division multiplexing, WDM）网络和弹性光网络（elastic optical network, EON）资源调度算法已无法继续满足需求。在这一背景下，面向多波段光网络场景的资源分配问题需要在考虑复杂物理效应的前提下进行研究。

为支持光网络由单波段向多波段演进，文献[7]提出在混合单波段/多波段弹性光网络中的需求配置问题，并设计了一系列启发式算法来解决路由、波段和频谱分配（route, wavelength, and band assignment, RWBA）问题。针对多波段扩容以及复杂物理效应引起的计算复杂度提升使传统启发式算法计算时间延长，文献[8]提出一种基于深度强化学习（deep reinforcement learning, DRL）的资源分配策略来缩短计算时间，并通过模拟实验验证了其在降低阻塞概率（blocking probability, BP）和提高网络吞吐量方面的有效性。文献[9]则设计了一种列生成（column generation, CG）方法，以解决传统方法扩展性差的

问题。文献[10]融入质量传输 (quality of transmission, QoT) 感知, 采用多周期规划和低裕度 (low margin) 的预测总通道数算法, 实现了不需要重新配置先前服务请求的目标, 进而缩短了资源分配的执行时间。为进一步提升单链路传输能力, 文献[11]结合空分复用技术和多波段技术, 利用整数线性规划模型解决了 RWBA 问题。

### (2) 多路径光层资源分配技术

仅靠单一链路的扩容难以满足智算业务的高实时性需求, 面对大粒度业务需求, 多路径光层资源分配技术可以通过构建并行光传输通道, 进一步提升传输速度, 以响应智算业务的高实时传输需求。多路径光层资源分配技术的核心在于突破传统单路径路由的带宽瓶颈, 在波长连续性约束和频谱碎片限制下实现多维度的资源协同。

文献[12]提出了一种基于人工神经网络 (artificial neural network, ANN) 辅助的调度方法, 用于融合了计算、存储和网络资源的光计算能力网络 (optical computing power network, OCPN) 中的大量数据传输。文献[13]通过拆分数据、采用多条路径传输以及临时存储技术, 提升了端到端时延敏感传输的带宽利用率, 缩短了交付时间, 并加速了数据传输过程, 实现了结合数据中心存储和多路径路由的联合调度。值得注意的是, 为保障业务的实时传输, 多路径的传输间应当控制传输时延差。当到达智算中心后, 多路径传输通常还需要借助智算中心内部的数据包重排机制, 辅助实现智算业务的同步。

### (3) IP/光跨层流量疏导技术

跨层资源分配技术可以显著降低网络带宽的浪费, 其将多个源自相同节点并行传输的业务有效整合到单一的光路中, 从而在一定程度上节省了网络资源, 并降低了网络端口的使用率。流量疏导技术则根据不同的业务需求 (如带宽、时延、优先级等) 进行合理的资源调度, 为高实时任务的优先调度和低时延传输提供了保障。

以 IP-over-WDM 为承载架构的跨层资源分配研究已取得了一定进展。文献[14]提出基于虚拟传输链路的带宽实时调整 (virtual transport link-dynamic bandwidth adjustment, VTL-DBA) 机制, 可以对端到端 (end-to-end, E2E) 业务进行灵活管理, 从而提高了网络资源的利用率。文献[15]使用已有的通用多协议标签交换 (generalized multi-protocol label switching, GMPLS) 技术和波长交换光网络 (wavelength switched optical network, WSON) 技术, 在 WDM 网络中建立通信机制, 使光层能够获取 IP 层的实际流量信息, 并根据实际流量采用动态调整服务等级协议 (service level agreement, SLA) 的方法来弹性分配资源。

WDM 网络通常采用固定频率槽和单一速率传输方式, 这限制了其在动态流量环境中的灵活性, 因此文献[16]提出基于 IP/MPLS 的 EON 分布式路由方案和连接建立机制, 探讨了 IP-over-EON 的分层流量资源分配问题。IP 层利用流量工程 (traffic engineering, TE) 动态选择流量路径, 并基于流量优先级和服务质量 (quality of service, QoS) 需求进行调整, 而光层则通过可切片带宽可变转发器 (sliceable bandwidth-variable transponder, SBVT) 和灵活的频谱分配策略动态建立弹性光路径, 以适应动态变化的流量需求, 提高网络资源的利用率。文献[17]通过研究部分资源卸载, 提出了一种新的整数线性规划 (integer linear programming, ILP) 模型和多种启发式方法, 有效优化了网络资源分配并降低了 E2E 时延。

## 2.2 高实时调度技术

超大规模组网场景下的资源分配效率问题成为制约智算中心互联的关键瓶颈。当网络规模扩展至数百节点、数千波长量级时, 传统 ILP 方法的计算耗时呈现指数级增长, 难以满足智算业务的高实时调度需求。

当前技术突破主要沿着两个方向展开: 一是



基于图神经网络拓扑特征提取的技术，通过将大规模网络映射至低维嵌入空间，将复杂优化问题转化为特征空间的高效搜索，实验证明，在大规模网络中，该方法仍能在计算时间上优于传统的启发式方法<sup>[18]</sup>；二是可行域压缩技术，通过注入业务流量模式先验知识，结合模拟退火与AI智能决策的混合优化框架，可大幅降低弹性光网络资源分配的计算复杂度<sup>[19]</sup>。分布式协同架构的创新同样值得关注，联邦学习框架通过参数共享机制实现了多域资源联合优化，在确保数据隐私性的前提下，使跨域业务建立成功率提升30%以上<sup>[20]</sup>。两类高实时调度技术如图2所示。

### 3 智算光网络自适应优化技术

为应对高突发智算业务带来的拥塞问题与资源利用率问题，新兴技术利用网络侧和端侧技术，协同实现前瞻性网络优化。一方面，在网络侧利用光层资源重整技术以及跨层网络重构技术来提升网络资源利用率。自适应网络优化技术将被动调整转为主动适应，打破了网络静态结构，利用动态化网络适配动态业务需求。另一方面，利用端侧协同的长距RDMA流量控制技术以及光传送网（optical transport network, OTN）流量控制技术，配合光层的流量调度技术，可以实现长距离无损传输，从而解决拥塞问题。

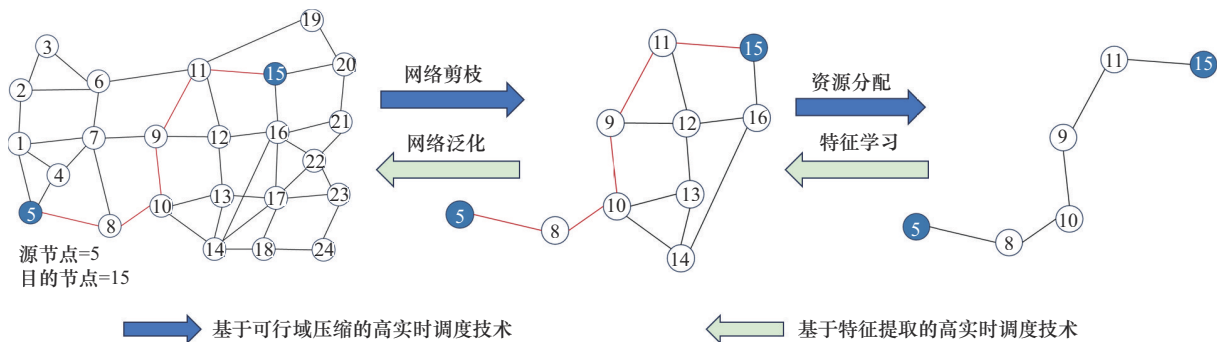


图2 两类高实时调度技术

### 3.1 网络侧自适应资源优化技术

业务的突发性使得网络中存在大量的碎片化空闲资源，为进一步提升网络资源利用率，降低网络拥塞概率，需要采用网络侧自适应资源优化技术，具体可以包括业务适配网络的光层资源重整技术以及网络适配业务的跨层网络重构技术。对于光层资源重整技术，考虑多波段的引入带来的额外物理效应，又包括频谱碎片整理与带宽碎片整理；对于跨层网络重构技术，主要根据业务需求对网络拓扑进行最优化估计，并根据新拓扑进行网络的重新规划。通过业务适配网络与网络适配业务的双向配合，网络侧自适应资源优化技术可以有效解决由智算业务产生的突发性新问题。

#### (1) 光层资源重整技术

智算业务的突发性使得光网络中存在频谱碎片与带宽碎片，当碎片程度达到一定阈值时，通过弹性光网络中的频谱碎片整理和多波段光网络中的带宽碎片整理，网络能够灵活应对需求变化，提升光网络资源利用效率。随着智能化和自动化算法的发展，碎片整理技术将为智算中心间的高效互联提供更强大的技术支撑，推动光网络向更高效、可持续的方向发展。

#### 1) 弹性光网络中的频谱碎片整理技术

光网络资源分配中的频谱一致性限制与频谱连续性限制使得网络中产生了频谱碎片，进而限制了网络容量<sup>[21]</sup>。频谱碎片整理是指对已分配的

网络资源进行重整,将分散的、无法使用的频谱资源重新聚合在一起,从而减轻光网络中的频谱碎片程度,增大网络容量。光网络中的频谱碎片整理问题通常被建模为ILP问题<sup>[22]</sup>,并被分为两个子问题:何时进行频谱碎片整理与如何进行频谱碎片整理。

根据整理触发时刻的不同,将频谱碎片整理分为反应式频谱碎片整理与主动式频谱碎片整理。反应式频谱碎片整理通常在新业务到达网络时触发;主动式频谱碎片整理则不需要等待新业务到达网络。例如,文献[23]在动态网络资源分配过程中,采用每间隔提前设定的时间主动进行一次频谱碎片整理。

在确定整理触发时刻后,即可采用不同的算法进行网络资源重整。文献[24]提出了一组启发式算法对已建立的业务连接进行重新分配,将研究人员观察到的网络现象通过一系列规则应用到频谱碎片整理中。此外,也有许多研究人员将人工智能算法应用到频谱碎片整理中来捕捉更深层次的网络运行特点,从而优化频谱碎片整理。文献[25]引入了一种基于无监督机器学习技术的新型碎片整理方法,通过对光路进行聚类来重新排列碎片化的光谱。文献[26]提出了用于频谱碎片整理的深度强化学习框架DeepDefrag,使用深度强化学习对网络资源进行重新分配。

## 2) 多波段光网络中的带宽碎片整理技术

为了满足智算中心等大流量业务的传输需求,现有网络正在从单波段向多波段光网络演进。多波段的引入使得波段间的非线性效应变得不可忽视,且由于业务的高突发性,信道的传输质量随之发生变化,当信道能够支持的传输质量超过业务的实际传输需求时,可以考虑提升调制格式以优化传输效率。由于WDM信道带宽一定,调制格式越高,可传输的比特率就越高。因此,当信道质量变好时,调制格式升级,可传输比特率提升,信道中就会存在剩余带宽,从而产

生带宽碎片。

带宽碎片整理方法主要分为启发式和机器学习两种。文献[27]提出了一种高效估算并调整业务对其他业务信噪比影响的方式,并基于此设计了一种序列分配业务的启发式算法。文献[28]提出了一种基于光信噪比的自适应资源分配算法,该算法在频谱分配时优先选择信噪比最优的频段,以最大化网络容量。此外,基于深度强化学习的带宽碎片整理方法逐渐受到关注。文献[29]提出了一种传输质量感知的光网络资源分配方案,该方案通过在业务调整前预估传输质量,旨在优化传输质量并最小化阻塞率。文献[30]进一步提出了利用深度强化学习进行业务分配与调度的方案,并设计了创新性的状态表示与奖励函数,以捕捉业务传输质量的动态变化。实验结果表明,该方案在减少阻塞率和提高资源分配效率方面显著优于传统方案。

## (2) 跨层网络重构技术

跨层网络重构技术是实现网络与业务需求匹配、优化资源调度的关键技术。在传统静态网络中,节点间的连接关系和链路容量是静态配置的,这意味着网络资源的分布是固定的。跨层网络重构技术以光层为传输底座来调整上层的IP网络,利用网络动态性适配业务的动态性,解决高突发性业务带来的网络问题。结合流量预测等人工智能技术,可以进一步提升技术的自适应性,为网络服务提供有力支撑。对于大模型训练,由于训练存在持续时间,频繁网络重构可能会带来业务动荡,因此,网络重构需要对范围和频率进行规划。考虑智算中心间的协同往往存在集群特征,在实际情况下,可以当网络碎片化达到一定程度时再进行局部的网络重构,以降低对业务的影响。

文献[31-32]基于深度神经网络设计了一种最优IP网络拓扑估计方法,并在此基础上设计了一种动态网络重构机制,反转控制方向,运用灵活光



网络的重构能力，由底层传输网络主动地根据业务带宽需求动态重构IP网络拓扑及链路容量，调节网络资源分布以适配需求分布，实现按需供给，提高了网络资源的利用率。

### 3.2 端侧协同的流量控制技术

智算业务的突发可能会导致网络的拥塞，致使业务性能下降。端侧协同的流量控制技术可以精准调整传输策略。在发生网络拥塞时，为实现流量控制，长距RDMA流量控制技术、OTN流量控制技术和光层流量调度技术之间相互依赖、协同工作。长距RDMA流量控制需要OTN流量控制的支持来实现长距离无损传输，而OTN流量控制则依赖光层的波长级调度和保护机制来实现高效的流量调度。光层为OTN流量控制提供高质量的传输通道，并将监测信息反馈给OTN流量控制模块，以实现更精准的流量优化。因此，长距RDMA流量控制技术和OTN流量控制技术共同提供流量控制的解决方案，为自适应网络优化提供有力技术支撑。端侧协同的远距离RDMA-OTN流量控制技术如图3所示。

#### (1) 长距RDMA流量控制技术

RDMA的卓越性能高度依赖底层的无损传输特性。与传统的基于内核的传输控制协议/网际协议（transmission control protocol/Internet protocol, TCP/IP）相比，RDMA网卡上的有限内存资源难以支持TCP中复杂的快速重传或选择性重传机制。因此，对于RDMA的流量控制研究是保障智算中心网络性能优越的关键。

常见的长距RDMA包括无损和有损两个方向。无损RDMA，依赖于优先级流量控制（priority flow control, PFC）机制，需要深度缓冲区来吸收飞行中的数据包，导致长距离链路的吞吐量下降。有损RDMA，如改进的基于以太网协议的RDMA（RDMA over converged Ethernet, RoCE）技术，虽然避免了PFC的头阻塞问题，但固定大小的位图设计无法同时满足不同往返时间（round-trip time, RTT）的流量性能需求。同时，长距离链路的高时延和大带宽时延积（bandwidth-delay product, BDP）进一步加剧了现有RDMA解决方案的性能瓶颈。

在此背景下，文献[33]提出了SWING方案，通过在长距离链路两端部署“中继”设备，频繁生成循环PFC信息，像“摆动”一样使开关队列长度在两阈值附近波动，将PFC信号在本地和远程交换机之间进行透明转发，从而降低了长距离链路的PFC响应时延。同时，中继设备通过周期性地生成高频PFC信号，强制本地设备以远程交换机的消耗速率发送数据，从而避免了额外缓冲区的需求。针对PFC存在的局限性，文献[34-35]提出了Bifrost，一种面向长距离跨数据中心链路的下游驱动无损流量控制机制。Bifrost通过维护一个“虚拟飞行数据包”的历史记录来预测未来到达的数据量，从而提前控制上游的发送速率。文献[36]提出了链路分段拥塞控制（link-segmented congestion control, LSCC）算法，该算法通过在数据中心的出口交换机上部

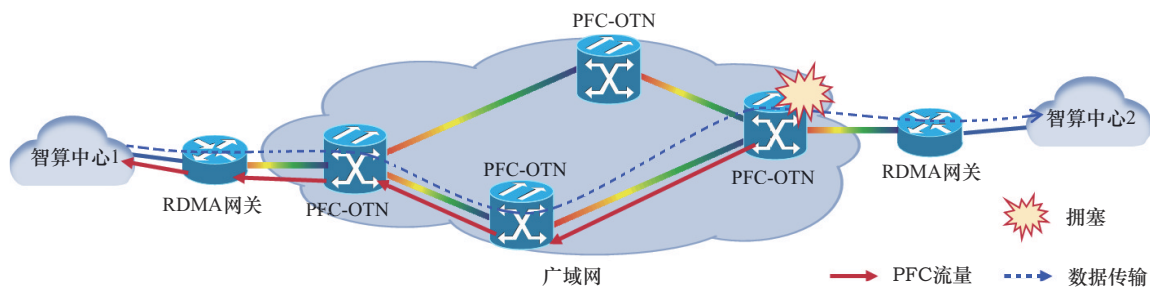


图3 端侧协同的远距离RDMA-OTN流量控制技术

署速率计算模块和速率限制模块，建立了一个分割的反馈环路，不仅为流量源提供了及时的拥塞反馈，还通过反馈机制显著减少了交换机的缓冲区压力。文献[37]提出了一种名为LoWAR (lossy wide-area RDMA) 的技术，旨在通过透明的前向纠错 (forward error correction, FEC) 技术提升RDMA在易失性广域网 (WAN) 中的性能。LoWAR将RDMA消息作为保护目标，而不是传统的数据报或字节流。它采用“包对齐编码模型”，为每个连接的RDMA消息分配编码块，并根据消息长度动态调整编码块的大小。这种方法避免了跨消息编码的低效性，并通过“基于缓冲区的更新式计算”降低了对缓存的需求。

#### (2) OTN流量控制技术

长距RDMA流量控制需要OTN流量控制的支持来实现长距离无损传输。OTN的流量控制机制可以为RDMA流量提供稳定的传输通道，并在拥塞发生时进行有效的流量调度。

通过与交换机、服务器等网络设备进行端网协同，OTN可以辅助实现拥塞控制。这种协同机制可以实时监测网络状态，并根据网络的拥塞情况动态调整数据传输速率，从而避免网络拥塞的发生，提高网络的利用率和传输效率。文献[38]指出，长距离传输网络设备需要响应流量控制信号，OTN传输设备必须具备PFC能力，能够缓存网络上的流量以确保无丢包，并能够逐步向上游推送PFC流量，从而与DCN交换机协作实现长距离无损传输。文献[38]通过OTN、端侧和RoCE交换机之间的跨域协作，提升了RDMA长距离传输能力。OTN设备实时感知PFC信息，当检测到拥塞时，立即在近端OTN端口限制流量以减少智算中心网络流量，同时将回压信息向上游传输，通过重用OTN设备的客户端缓冲资源，持续缓冲链路流量，直到源端回压成功。实验证明，经过PFC-OTN后，端口带宽利用率从20%

提升至超过90%，并实现了TB级海量数据在1 min内的传输。

## 4 智算光网络前摄性故障恢复技术

面对智算网络承载业务的复杂度与实时性需求，传统被动式故障恢复机制因检测时延和恢复效率已显不足，难以满足高可靠场景的需求。现代网络正加速向“实时检测—预测定位—快速恢复”的智能化方向演进：一方面，通过物理层光性能监测与网络层协议探针构建多维度感知体系，突破单点检测瓶颈；另一方面，融合机器学习与深度学习算法，从历史数据中挖掘潜在故障模式，实现故障前干预，最终结合量子计算加速的资源调度与IP/光协同重构技术，形成业务级与节点级双重保障的快速恢复能力。这一技术演进框架为智算光网络构建起“感知—决策—执行”的从被动性到前摄性的闭环防御体系。

### 4.1 分层故障检测技术

故障检测的速度和准确性是影响网络可靠性的重要因素。故障检测技术通过分层协同机制来保障网络的高可用性，其核心涵盖物理传输介质的基础性能监测与网络架构层的动态状态感知两大维度。基于物理传输性能检测与网络层状态感知的分层故障检测技术如图4所示，在物理层，光信号传输质量直接影响网络承载能力，需要通过多维参数构建精准的底层健康评估体系；在网络层，链路连通性与业务流完整性则依赖协议级探针实现全局状态追踪。二者分别从光域特征解析与逻辑拓扑验证的角度切入，共同突破传统单点检测的局限性，为智算光网络提供从物理损伤定位到业务层异常感知的全栈故障诊断能力。

#### (1) 物理层光性能检测技术

光性能监测 (optical performance monitoring, OPM) 技术可以为光性能检测实时收集性能指标，是光纤通信网络中保障系统稳定性和可靠性的关键手段之一<sup>[39]</sup>。传统的技术通过实时监测光

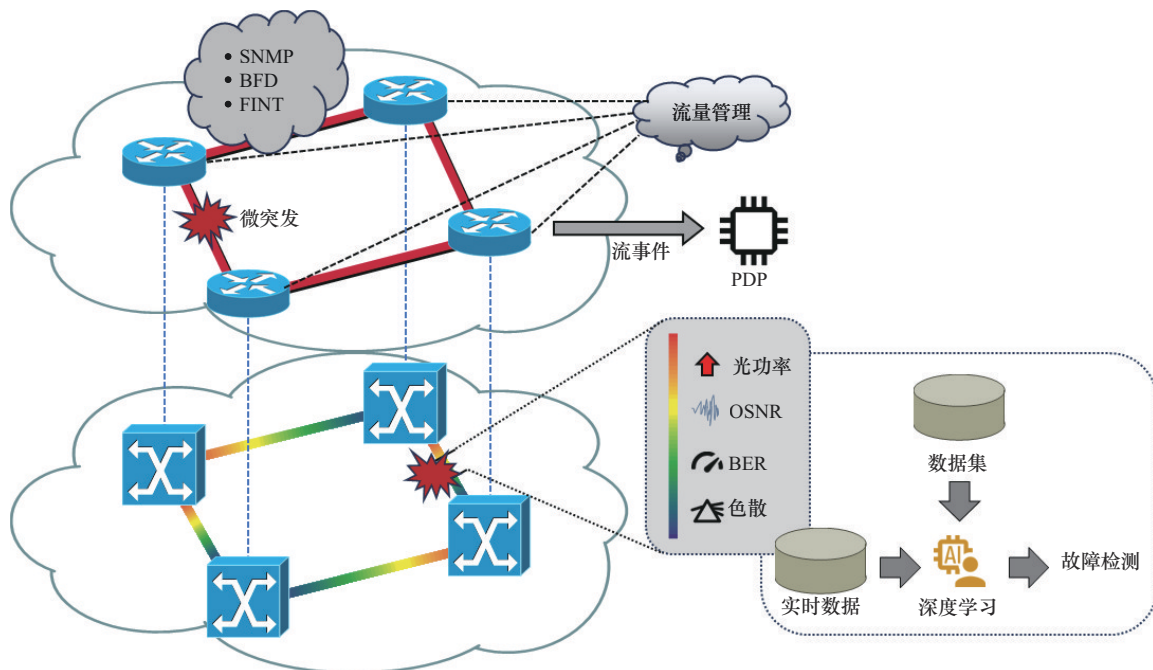


图4 基于物理传输性能检测与网络层状态感知的分层故障检测技术

纤网络的物理层多维度参数<sup>[40]</sup>，包含光功率、光信号噪声比（optical signal-to-noise ratio, OSNR）<sup>[41]</sup>、误码率（bit error ratio, BER）、Q因子、色散<sup>[42-43]</sup>及偏振模色散等，构建起光网络传输质量的健康评估体系。其中，光功率直接反映信号的传输强度，OSNR精准量化信号的质量，BER则表征数据传输的可靠性。通过实时监控物理层的光纤传输质量，光性能监测不仅为网络维护提供了必要的诊断工具，还能在网络出现潜在问题时发出预警，从而避免系统故障的发生。

QoT估计在光网络中是一个复杂问题，但其准确性至关重要。传统的物理模型在处理QoT估计问题时存在模型不准确、速度慢和依赖不确定参数等问题，因此，研究人员将机器学习与现有的网络光信号处理技术结合<sup>[44-45]</sup>，形成了QoT估计的系列解决方案。现有的与机器学习相关解决方案主要采用监督、半监督或无监督机器学习（machine learning, ML）方法，文献[46]提出了一种混合学习解决方案，结合监督、无监督和协

作式ML的优点，利用基于密度的含噪声空间聚类应用（density-based spatial clustering of applications with noise, DBSCAN）方法进行模式分析，标记异常点为潜在故障，采用自学习机制训练神经网络分类器实现在线故障检测。然而，训练这些基于机器学习的故障检测算法通常需要大量已标记的故障样本，而实际系统中故障样本稀少，文献[47]针对此进一步提出了一种基于生成对抗网络（generative adversarial network, GAN）的故障检测和识别算法，编码器、解码器和判别器的对抗，使得训练该识别算法所需的故障样本数量大幅减少。

## （2）网络层性能检测技术

在传统网络中，监控网络性能和故障常依赖于简单网络管理协议（simple network management protocol, SNMP）、端口计数器、系统日志（syslogs）以及毫秒级的双向转发检测（bidirectional forwarding detection, BFD）机制。这些工具可以在秒级时间内检测到链路故障和网络设备的响应问题。例如，SNMP提供了对网络设备状态的简

单查询,允许网络管理员监控设备的运行状况;端口计数器和syslogs则能够帮助收集网络流量数据和系统事件,以分析潜在的故障。

然而,这些机制通常只能检测到较为简单的网络问题,如设备无响应或者端口故障,存在粒度粗、开销大、无法精准定位等问题。Net-Seer<sup>[48]</sup>是一种基于流事件遥测(flow event telemetry, FET)的网络监视器,它利用可编程数据平面(programmable data plane, PDP)实现对流的持续监测,同时依靠专用集成电路(application-specific integrated circuit, ASIC)、CPU和RAM资源协作,确保无漏报并高效消除误报。FINT<sup>[49]</sup>是一种灵活的带内网络遥测方法,通过设计三重位图机制和灵活的带内网络遥测(in-band network telemetry, INT)字段结构,同时结合贪心算法选择遥测元数据,在不显著增加带宽消耗的情况下,有效降低了携带INT数据的小流的平均流完成时间,实现了对数据中心网络微突发的遥测,满足了对网络状态及时感知的需求。F4Tele<sup>[50]</sup>利用自由空间光(free space optical, FSO)通信和WDM构建了一个灵活且高性能的逻辑网络,该逻辑网络专门用于管理流量的传输,在不影响数据网络的情况下传输网络管理流量(network management traffic, NMT),进而保证了用于故障检测的NMT数据的完整性和时效性。

#### 4.2 前摄性故障定位技术

传统的故障检测与恢复的方法往往只能在故障发生后进行被动响应,在许多情况下已经无法满足高效、可靠的网络运营需求。因此,为了实现前摄式故障检测恢复,前摄性故障定位技术为其提供了前所未有的保障。前摄性故障定位技术的核心在于提前识别潜在故障,并通过实时监测和分析将这些故障预测出来。

将机器学习和深度学习算法应用于网络的监控数据中,能够实时捕捉到信号中的异常变化,并根据历史故障数据进行模式识别,预测出可能

出现的故障。这不仅减少了传统故障恢复所需的时间,还能够在故障发生之前及时采取措施,避免出现大规模的服务中断和资源浪费。

文献[51]提出了一种基于多层双向长短期记忆(bidirectional long short-term memory, Bi-LSTM)网络的故障预测算法。该算法通过处理时间序列数据,能够捕捉数据中的长期依赖关系,并且能够同时考虑更接近和更远离预测时间点的特征,从而更准确地确定数据的权重。文献[52-53]提出了时序数据驱动的光网络设备故障智能预测方案,先基于长短期记忆(long short-term memory, LSTM)算法进行故障预测建模,利用OTN单板的历史运行数据,可提前一天预测其运行状态,同时将点乘注意力机制嵌入模型,可识别OTN单板潜在故障原因并提高故障预测性能。文献[54]以决策树为基础构建了一种机器学习模型,该模型能够依据设备当前状态预测其在未来特定时间段内是否会发出警报,同时,将光网络设备故障预测问题建模为分类问题,通过分析输入数据的特征,对故障情况进行分类判断。

#### 4.3 快速故障恢复技术

为实现智算中心间互联网络的高可靠性,在检测出网络故障后,如何快速恢复网络连通性显得至关重要。光网络在智算中心间互联中的故障恢复技术分为保护倒换式和非保护倒换式。保护倒换式通过预设路径实现快速切换,但会浪费带宽并增加运维成本;非保护倒换式通过动态计算实时寻找替代路径,提高了资源利用效率,更适合复杂的智算中心间互联网络。在此基础上,从业务级和节点级两个层面考虑快速故障恢复。业务级快速故障恢复技术侧重于在业务层面上实现故障的快速检测和恢复,如通过多种重路由方式来减少故障对用户服务的影响。节点级快速故障恢复技术则通过节点的快速恢复对多业务故障进行并行处理,进一步增强了网络的韧性,缩短了



故障恢复时间，确保了智算中心间互联网络的持续稳定运行。

### (1) 业务级快速故障恢复技术

业务级快速故障恢复技术主要针对网络服务层面的故障恢复，目的是最小化网络故障对终端用户服务的影响。该技术的核心在于能够在业务传输过程中，快速识别和响应网络中的故障情况，从而实现业务的持续性和服务的高可用性。

目前常用的解决方案是重路由方案，其作为一种有效的被动保护机制，旨在确保网络故障发生后业务不中断。由于数据平面的处理时延通常比控制平面低几个数量级，现代网络通常会采用数据平面内的快速重路由（fast re-route, FRR）机制。在该机制下，当故障发生时，网络中的数据流会被快速地引导至新路径，从而保证业务的持续性，缩短故障恢复的时间。文献[55]提出一种名为拓扑无关的无环备用路径（topology independent loop-free alternate, TI-LFA）的快速重路由技术，并将其与段路由相结合。段路由<sup>[56]</sup>是基于源路由的网络技术，旨在简化IP网络中的流量工程和网络管理。段路由通过在数据包头部插入有序的指令列表来指导数据包在网络中的转发路径。基于IPv6转发平面的段路由（segment routing over IPv6, SRv6）特别适用于支持网络编程模型，能够实现更复杂的服务，支持智算中心间互联网络业务。文献[57]提出了动态链路风险重路由算法，通过建立业务重要度与链路风险评估模型，优先为高重要度的故障业务分配低风险的恢复资源，保证了重要业务的可靠传输。

故障业务恢复时间越短，则代表网络对故障的处理能力越强，生存性能越好。因此，为保证发生故障后智算中心间互联网络的稳定性以及业务的服务质量，在保证可靠传输的同时，研究更快速的故障恢复算法更为迫切。相较于经典计算机，专用量子计算机（如测量反馈相干伊辛机、量子退火机）具有更快速解算组合优化问题的优

点，可以利用量子力学原理同时执行多个计算，为诸多NP难组合优化问题提供更快速、更精确的解决方案<sup>[58]</sup>。由于量子技术发展尚处于嘈杂的中尺度量子（noisy intermediate scale quantum, NISQ）时代<sup>[59]</sup>，混合量子-经典计算是解决NISQ时代计算机局限性的有效方法<sup>[60]</sup>，因此，在经典计算方法基础上，引入量子计算技术，将是提升业务级故障恢复速度的强有力措施。

### (2) 节点级快速故障恢复技术

网络中单个节点或多个节点发生故障时可能会造成多业务并发故障，在智算中心互联场景中，多业务并发故障的恢复效率直接决定了网络韧性。传统串行恢复机制中资源竞争和时序冲突易引发恢复时延，同时，恢复时间与业务数量成正比，当业务数量较大时业务级故障恢复效率低，节点级快速故障恢复技术则可实现多故障域并行恢复。

针对网络中的单节点故障恢复，文献[61-62]提出一种节点级快速故障恢复方法，利用文献[63]所构建的跨层协同控制器，通过文献[59]所提出的双迁移的单节点恢复方法以及文献[61]所提出的多节点协同恢复方法，可以实现节点级故障恢复。这种节点级快速恢复技术不需要改变重构前后的逻辑视图，因此，可以实现大量业务的同时恢复，进而减少故障恢复时间。同时，这种方法利用光网络的灵活性为业务传输提供了额外的带宽资源，可以有效提升故障恢复成功率。

针对跨区域链式故障或级联故障场景，节点级恢复技术通常采用分层递进策略。文献[64]综合考虑了网络状态、故障节点的度分布和中心性等因素，对节点进行评分，对得分高于特定阈值的故障节点进行优先恢复，随后采用增量式恢复机制，以核心节点为锚点逐跳扩展至边缘节点。

## 5 未来研究展望

随着智算光网络向着网络资源复杂化、网络

规模扩大化发展,如何更有针对性地满足智算业务与智算光网络对高实时调度与深度融合的需求,成为一个重要课题。对此,研究人员可以从高实时大规模调度技术、算网深度协同技术以及智算数字孪生技术这3个角度出发,对未来技术的发展进行展望。

### 5.1 高实时大规模调度技术

智算业务的高实时性需求要求其在大规模场景下实现高效调度。同时,算网融合导致多维资源管理的复杂性显著增加,引发了“维度爆炸”问题。然而,当前技术演进仍面临多重挑战。例如,在线学习场景下的样本效率不足问题,导致强化学习算法收敛速度滞后于业务调度需求;物理层损伤模型的精度缺陷可能引发虚拟拓扑性能预测偏差;多目标优化中的求解效率制约方案优选速度等。因此,开发高实时调度方法成为当务之急。新型高实时调度方法需要能够应对大规模网络环境,并能够有效处理多维度问题,以确保智算业务的高效运行和资源的合理利用。

新型高实时调度方法可以引入复杂网络技术,通过对网络行为的建模并结合流量预测方法,对网络状态进行分析,提前进行资源分配和调度;也可以提取网络特征,通过在不同维度对网络特征的剪枝,共同实现高效的资源分配;还可以引入新兴的快速计算方法,如随着量子计算的进一步发展,可以从混合量子-经典计算逐步变为纯量子计算,以提升计算效率。

### 5.2 算网深度协同技术

智算业务已经发展为从细粒度算力需求(如AI推理服务)到粗粒度算力需求(如大规模训练任务)的多层次业务。为了满足这些需求,智算中心需要实现虚拟图形处理单元(virtual graphics processing unit, vGPU)级到整机柜级的多粒度资源供给,目前,已经实现了异构资源的算力重组,未来有望将计算资源抽象为最小可调度单元(如1 TFLOPS),实现算力的动态重组,从而

满足智算业务的不同粒度算力需求。同时,智算中心要能够针对不同AI业务特征(如CV/NLP/科学计算)设计差异化调度策略,对于毫秒级响应的实时推理业务需求,可部署就近节点;对于高算力需求业务,可以实现跨地域分布式GPU集群协同,同时保障长时稳定的带宽供给。例如,在“东数西算”战略工程的背景下,将我国东部万亿参数大模型训练任务拆解后,动态分配至西部的多个智算中心。同时,在我国西部的算力节点内部,通过算网协同可以实现vGPU级资源精准供给。最终,这些努力将共同构建“算力调度全国一张网”,实现跨枢纽的算力秒级调度与交易。

### 5.3 智算数字孪生技术

数字孪生(digital twin, DT)技术作为一种新兴的技术范式,在智算中心间光互连网络中应用前景广阔,通过精确的物理链路建模和实时网络检测,能够提升网络的智能化水平,确保网络高效稳定运行。光网络作为智算中心内外部高速数据传输的关键基础设施,面临更高的带宽、更低的时延和更强的可靠性要求,需要更加智能化、高效化的管理和优化手段。DT技术通过数据采集和遥测技术,能够构建精准的物理链路模型,实时反映网络中光链路、设备节点等元素的状态,为智算中心间光互连网络提供持续、精准的监控。通过实时监测和仿真,DT技术还可以模拟未来可能出现的网络行为,帮助预测潜在的瓶颈或故障点。结合AI算法,DT技术能够自动检测网络异常,迅速定位问题,并实时调整网络配置、优化链路和设备参数,从而有效提高运维效率。

## 6 结束语

本文围绕智算中心间互联光网络面临的高实时性、高突发性和高可靠性挑战展开深入研究,提出一系列创新技术与解决方案,为智算光网络的发展提供了重要支撑。在资源分配方面,通过多波段光层资源分配、多路径光层资源分配以及



IP/光跨层流量疏导等技术, 实现全局规划的统筹协调, 能够满足智算业务的高实时传输与调度需求; 针对网络优化, 构建网络侧与端侧协同的自适应优化体系, 利用光网络资源重整、跨层网络重构以及端侧协同的流量控制技术, 有效应对智算业务高突发性带来的拥塞和资源闲置问题; 在故障恢复领域, 借助分层故障检测、前摄性故障定位和快速故障恢复技术, 实现主动干预的故障恢复, 保障了光网络的高可靠性。

展望未来, 高实时调度方法将借助复杂网络技术、流量预测和新兴计算方法, 突破多维资源管理难题; 算网深度协同技术将实现多粒度算力供给与差异化调度策略, 满足多样化智算业务需求; DT技术将通过精准建模和实时仿真, 提升光网络智能化管理水平。这些技术的融合发展, 将推动智算中心间互联光网络向更加智能、高效、可靠的方向迈进, 为人工智能产业的持续创新提供坚实的网络保障, 助力科技革命与产业变革不断深入。

## 参考文献:

- [1] 中国移动通信集团有限公司. 面向超万卡集群的新型智算技术白皮书[R]. 北京: 中国移动通信集团有限公司, 2024.  
China Mobile Communications Group Limited. White paper on new intelligent computing technology for over 10 000 card clusters[R]. Beijing: China Mobile Communications Group Co., Ltd., 2024.
- [2] 中国电信股份有限公司研究院. 分布式智算中心无损网络技术白皮书[R]. 北京: 中国电信股份有限公司研究院, 2024.  
China Telecom Research Institute. White paper on lossless network technology of distributed intelligence center[R]. Beijing: Research Institute of China Telecom Co., Ltd., 2024.
- [3] 浪潮信息, 中国信息通信研究院. 人工智能算力高质量发展评估体系报告[R]. 北京: 中国信息通信研究院, 2024.  
Inspur Information, China Academy of Information and Communications Technology. Report on high-quality development evaluation system of artificial intelligence computing power[R]. Beijing: China Academy of Information and Communications Technology, 2024.
- [4] CHOI W, SHIN M, LEE H, et al. Multi-task learning for real-time autonomous driving leveraging task-adaptive attention generator[C]//Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2024: 14732-14739.
- [5] 王光全, 满祥银, 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.  
WANG G Q, MAN X K, XU B H, et al. Deterministic optical transmission for wide area and long-distance computing power interconnection[J]. Designing Techniques of Posts and Telecommunications, 2024(2): 7-13.
- [6] SAMBO N, FERRARI A, NAPOLI A, et al. Provisioning in multi-band optical networks[J]. Journal of Lightwave Technology, 2020, 38(9): 2598-2605.
- [7] DIN D R. Heuristic algorithms for demand provisioning in hybrid single/multi-band elastic optical networks[C]//Proceedings of the 2023 IEEE 8th Optoelectronics Global Conference (OGC). Piscataway: IEEE Press, 2023: 8-12.
- [8] BENTERKIA, PEDRO J, EIRAA, et al. Deep reinforcement learning for resource allocation in multi-band optical networks[C]//Proceedings of the 2024 International Conference on Optical Network Design and Modeling (ONDM). Piscataway: IEEE Press, 2024: 1-4.
- [9] CHEN C, XIAO S L, ZHOU F, et al. Throughput maximization in multi-band optical networks with column generation[C]//Proceedings of the ICC 2024 - IEEE International Conference on Communications. Piscataway: IEEE Press, 2024: 3034-3039.
- [10] MEHRABI M, BEYRANVAND H, EMADI M J, et al. Efficient statistical QoT-aware resource allocation in EONs over the C+L-band: a multi-period and low-margin perspective[J]. Journal of Optical Communications and Networking, 2024, 16(5): 577-592.
- [11] WU Q, WANG J D, CHEN S B, et al. Resource allocation problem in multi-band space-division multiplexing elastic optical networks[C]//Proceedings of the 2022 18th International Conference on Computational Intelligence and Security (CIS). Piscataway: IEEE Press, 2022: 225-228.
- [12] LIN X, LIN H H, ZHANG C X, et al. ANN-assisted scheduling method for bulk data transfers in optical computing power networks[C]//Proceedings of the 2023 Opto-Electronics and Communications Conference (OECC). Piscataway: IEEE Press, 2023: 1-6.
- [13] LIN X, JI S, YUE S N, et al. Adaptive multi-path SnF scheduling method for delay-sensitive transfers across inter-datacenter optical networks[C]//Proceedings of the 2022 International Conference on Optical Network Design and Modeling (ONDM). Piscataway: IEEE Press, 2022: 1-6.

- [14] ZHOU Y, RAMAMURTHY B, GUO B L, et al. Supporting dynamic bandwidth adjustment based on virtual transport link in software-defined IP over optical networks[J]. *Journal of Optical Communications and Networking*, 2018, 10(3): 125-137.
- [15] SOOD N, KUMAR P. Determination of traffic utilization for DWDM networks and enhanced traffic engineering applications[J]. *Technical Disclosure Commons*, 2022: 1-15.
- [16] HALABI W. Distributed routing scheme for IP-based elastic optical networks (EONs)[C]//*Proceedings of the 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*. Piscataway: IEEE Press, 2023: 1-4.
- [17] CHEN B W, LIU L, FAN Y X, et al. Low-latency partial resource offloading in cloud-edge elastic optical networks[J]. *Journal of Optical Communications and Networking*, 2024, 16(2): 142-158.
- [18] ZHOU Z X, GU R T, ZHANG X Y, et al. Opti-DeepRoute: a topology-adaptive deep reinforcement learning based service provisioning framework for elastic optical network[C]//*Proceedings of the IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. Piscataway: IEEE Press, 2024: 1-2.
- [19] ANDREOLETTI D, ROTTONDI C, BIANCO A, et al. A machine learning framework for scalable routing and wavelength assignment in large optical networks[C]//*Proceedings of the 2021 Optical Fiber Communications Conference and Exhibition (OFC)*. Piscataway: IEEE Press, 2021: 1-3.
- [20] ZHANG X Y, GU R T, DONG J S, et al. Field trial of privacy-preserving resource allocation in multi-domain optical networks based on federated reinforcement learning[C]//*Proceedings of the 2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM)*. Piscataway: IEEE Press, 2023: 1-4.
- [21] CHATTERJEE B C, BA S, OKI E. Fragmentation problems and management approaches in elastic optical networks: a survey[J]. *IEEE Communications Surveys & Tutorials*, 2018, 20(1): 183-210.
- [22] ZHANG M Y, YOU C S, ZHU Z Q. On the parallelization of spectrum defragmentation reconfigurations in elastic optical networks[J]. *IEEE/ACM Transactions on Networking*, 2016, 24(5): 2819-2833.
- [23] CHATTERJEE B C, WADUD A, OKI E. Proactive fragmentation management scheme based on crosstalk-avoided batch processing for spectrally-spatially elastic optical networks[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(9): 2719-2733.
- [24] CALDERÓN F, LOZADA A, MORALES P, et al. Heuristic approaches for dynamic provisioning in multi-band elastic optical networks[J]. *IEEE Communications Letters*, 2022, 26(2): 379-383.
- [25] TRINDADE S, DA FONSECA N L S. Machine learning for spectrum defragmentation in space-division multiplexing elastic optical networks[J]. *IEEE Network*, 2021, 35(1): 326-332.
- [26] ETEZADI E, NATALINO C, DIAZ R, et al. DeepDefrag: a deep reinforcement learning framework for spectrum Defragmentation[C]//*Proceedings of the GLOBECOM 2022-2022 IEEE Global Communications Conference*. Piscataway: IEEE Press, 2022: 3694-3699.
- [27] MEHRABI M, BEYRANVAND H, EMADI M J. Multi-band elastic optical networks: inter-channel stimulated Raman scattering-aware routing, modulation level and spectrum assignment[J]. *Journal of Lightwave Technology*, 2021, 39(11): 3360-3370.
- [28] JANA R K, CHATTERJEE B C, SINGH A P, et al. Quality-aware resource provisioning for multiband elastic optical networks: a deep-learning-assisted approach[J]. *Journal of Optical Communications and Networking*, 2022, 14(11): 882-893.
- [29] TENG Y R, NATALINO C, ARPANA EI F, et al. DRL-assisted dynamic QoT-aware service provisioning in multi-band elastic optical networks[C]//*Proceedings of the 2024 European Conference on Optical Communication (ECOC)*. Piscataway: IEEE Press, 2024: 1-4.
- [30] ASIRIA, WANG B. Deep reinforcement learning for QoT-aware routing, modulation, and spectrum assignment in elastic optical networks[J]. *Journal of Lightwave Technology*, 2025, 43(1): 42-60.
- [31] LI Z K, GU R T, WANG L N, et al. Computing-aware proactive network reconfiguration for optical networks interconnected edge computing system[C]//*Proceedings of the 2021 Optical Fiber Communications Conference and Exhibition (OFC)*. Piscataway: IEEE Press, 2021: 1-3.
- [32] WANG L N, GU R T, LI Z K, et al. Computing-aware proactive IP-optical integrated network restructuring for edge computing[C]//*Proceedings of the 2021 19th International Conference on Optical Communications and Networks (ICOON)*. Piscataway: IEEE Press, 2021: 1-3.
- [33] CHEN Y Q, TIAN C, DONG J Q, et al. Swing: providing long-range lossless RDMA via PFC-relay[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2023, 34(1): 63-75.
- [34] YU P W, XUE F Y, TIAN C, et al. Bifrost: extending RoCE for long distance inter-DC links[C]//*Proceedings of the 2023 IEEE 31st International Conference on Network Protocols (ICNP)*.



- Piscataway: IEEE Press, 2023: 1-12.
- [35] HUANG C Y, XUE F Y, YU P W, et al. Minimizing buffer utilization for lossless inter-DC links[J]. *IEEE/ACM Transactions on Networking*, 2024, 32(6): 4960-4975.
- [36] LONG M F, HAN J P, WANG W T, et al. LSCC: link-segmented congestion control for RDMA in cross-datacenter networks[C]//*Proceedings of the 2024 IEEE/ACM 32nd International Symposium on Quality of Service (IWQoS)*. Piscataway: IEEE Press, 2024: 1-10.
- [37] ZUO T Y, SUN T, ZHU S Y, et al. LoWAR: enhancing RDMA over lossy WANs with transparent error correction[C]//*Proceedings of the 2024 IEEE/ACM 32nd International Symposium on Quality of Service (IWQoS)*. Piscataway: IEEE Press, 2024: 1-10.
- [38] TAN Y X, MAN X K, WANG G Q, et al. Field trial of long-distance RDMA lossless transmission for wide-area data center interconnection[C]//*Proceedings of the 2024 Asia Communications and Photonics Conference (ACP) and International Conference on Information Photonics and Optical Communications (IPOC)*. Piscataway: IEEE Press, 2024: 1-3.
- [39] ODA S, BOUDA M, GE Y, et al. Innovative optical networking by optical performance monitoring and learning process[C]//*Proceedings of the 2018 European Conference on Optical Communication (ECOC)*. Piscataway: IEEE Press, 2018: 1-3.
- [40] BARZEGAR S, RUIZ M, SGAMBELLURI A, et al. Soft-failure detection, localization, identification, and severity prediction by estimating QoS model input parameters[J]. *IEEE Transactions on Network and Service Management*, 2021, 18(3): 2627-2640.
- [41] CHENG Y J, ZHANG W K, FU S N, et al. Transfer learning simplified multi-task deep neural network for PDM-64QAM optical performance monitoring[J]. *Optics Express*, 2020, 28(5): 7607-7617.
- [42] WANG C X, FU S N, WU H, et al. Joint OSNR and CD monitoring in digital coherent receiver using long short-term memory neural network[J]. *Optics Express*, 2019, 27(5): 6936-6945.
- [43] FENG J C, JIANG L, YAN L S, et al. Intelligent optical performance monitoring based on intensity and differential-phase features for digital coherent receivers[J]. *Journal of Lightwave Technology*, 2022, 40(12): 3592-3601.
- [44] 李鸿, 刘武, 罗鸣. 基于机器学习的光网络监测与优化方法[J]. *光通信研究*, 2024(3): 5-14.
- LI H, LIU W, LUO M. Optical network monitoring and optimization methods based on machine learning[J]. *Study on Optical Communications*, 2024(3): 5-14.
- [45] POINTURIER Y. Machine learning techniques for quality of transmission estimation in optical networks[J]. *Journal of Optical Communications and Networking*, 2021, 13(4): B60-B71.
- [46] CHEN X L, LIU C Y, PROIETTI R, et al. Automating optical network fault management with machine learning[J]. *IEEE Communications Magazine*, 2022, 60(12): 88-94.
- [47] LUN H Z, LIU X M, CAI M, et al. GAN based soft failure detection and identification for long-haul coherent transmission systems[C]//*Proceedings of the Optical Fiber Communication Conference (OFC) 2021*. Optica Publishing Group, 2021: 1-3.
- [48] ZHOU Y, SUN C, LIU H H, et al. Flow event telemetry on programmable data plane[C]//*Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication*. New York: ACM, 2020: 76-89.
- [49] XIE S X, HU G Y, XING C Y, et al. FINT: flexible in-band network telemetry method for data center network[J]. *Computer Networks*, 2022, 216: 109232.
- [50] ALGHADHBAN A. F4Tele: FSO for data center network management and packet telemetry[J]. *Computer Networks*, 2021, 186: 107711.
- [51] GAO J C, WANG H Y, SHEN H Y. Task failure prediction in cloud data centers using deep learning[J]. *IEEE Transactions on Services Computing*, 2020, 15(3): 1411-1422.
- [52] 张春宇. 光网络系统故障智能预测与分析技术研究[D]. 北京: 北京邮电大学, 2022.
- ZHANG C Y. Research on intelligent fault prediction and analysis technology of optical network system[D]. Beijing: Beijing University of Posts and Telecommunications, 2022.
- [53] ZHANG C Y, WANG D S, JIA J W, et al. Attention mechanism-driven potential fault cause identification in optical networks[C]//*Proceedings of the 2021 Optical Fiber Communications Conference and Exhibition (OFC)*. Piscataway: IEEE Press, 2021: 1-3.
- [54] MURPHY K, LAVIGNOTTE A, LEPERS C. Fault prediction for optical access network equipment using decision tree methods[C]//*Proceedings of the 2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM)*. Piscataway: IEEE Press, 2023: 1-5.
- [55] ABHISHEK SINGH J, SACHIN K M R, SHUSHRUTHA K S. Implementation of topology independent loop free alternate with segment routing traffic[C]//*Proceedings of the 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. Piscataway: IEEE Press, 2021: 1-5.

- [56] VENTRE P L, SALSANO S, POLVERINI M, et al. Segment routing: a comprehensive survey of research activities, standardization efforts, and implementation results[J]. IEEE Communications Surveys & Tutorials, 2021, 23(1): 182-221.
- [57] 郭学让, 蒋一可, 李亚平, 等. 持续性灾害下风险感知的光网络业务恢复算法[J]. 光通信研究, 2024(2): 10-15.  
GUO X R, JIANG Y K, LI Y P, et al. Risk-aware optical network service restoration algorithm under persistent disasters[J]. Study on Optical Communications, 2024(2): 10-15.
- [58] KRAUSS T, MCCOLLUM J. Solving the network shortest path problem on a quantum annealer[J]. IEEE Transactions on Quantum Engineering, 2020, 1: 3101512.
- [59] BHARTI K, CERVERA-LIERTA A, KYAW T H, et al. Noisy intermediate-scale quantum algorithms[J]. Reviews of Modern Physics, 2022, 94(1): 015004.
- [60] OSABA E, VILLAR-RODRIGUEZ E, GOMEZ-TEJEDOR A, et al. Hybrid quantum solvers in production: how to succeed in the NISQ era?[J]. arXiv preprint, 2024: 2401.10302.
- [61] LIAN M, GU R T, QU Y Y, et al. Flexible optical network enabled hybrid recovery for edge network with reinforcement learning[C]//Proceedings of the 2020 Optical Fiber Communications Conference and Exhibition (OFC). Piscataway: IEEE Press, 2020: 1-3.
- [62] LI Y X, GU R T, JI Y F. Multi-node cooperative recovery against IP node failure enabled by flexible optical network[C]//Proceedings of the 2023 Optical Fiber Communications Conference and Exhibition (OFC). Piscataway: IEEE Press, 2023: 1-3.
- [63] GU R T, QU Y Y, LIAN M, et al. Flexible optical network enabled proactive cross-layer restructuring for 5G/B5G backhaul network with machine learning engine[C]//Proceedings of the 2020 Optical Fiber Communications Conference and Exhibition (OFC). Piscataway: IEEE Press, 2020: 1-3.
- [64] Su Q Y, Sun J X, LI J. Node recovery optimization of cyber-physical power systems based on the probability of fault propagation[C]//Proceedings of the 2024 China Automation Congress (CAC). Piscataway: IEEE Press, 2024: 140-145.

## [作者简介]



李韞瑄 (1998-), 女, 北京邮电大学信息与通信工程学院博士生, 主要研究方向为IP/光网络和智算中心光网络。



杨亚萍 (2002-), 女, 北京邮电大学信息与通信工程学院硕士生, 主要研究方向为智算中心光网络。



涂佳一 (2002-), 女, 北京邮电大学信息与通信工程学院硕士生, 主要研究方向为智算中心光网络。



顾仁涛 (1983-), 男, 北京邮电大学信息与通信工程学院教授、博士生导师, 主要研究方向为智能信息网络和智算光网络。



纪越峰 (1960-), 男, 北京邮电大学信息与通信工程学院教授、博士生导师, 北京邮电大学信息光子学与光通信全国重点实验室副主任, 主要研究方向为宽带通信网络和光通信技术。