面向新型智算中心的以太网弹性通道（FlexLane）技术白皮书

（2025 年）

发布单位：中国移动通信有限公司研究院

前　言

随着以 ChatGPT、Deepseek 为代表的 AI 大模型崛起，算力需求呈指数级增长，全球正加速建设智算中心以应对这一挑战。智算中心内部或智算中心间海量的数据交换，对网络链路的可靠性提出了前所未有的要求。任何链路闪断或中断都可能导致 AI 训练任务失败，造成巨大的时间和资源浪费。然而，光模块的成本与可靠性瓶颈以及大规模集群中链路数量的激增，使得已有技术难以满足新型智算中心 AI 业务对可靠性的需求。

本白皮书面向新型智算中心逐渐以承载 AI 业务为主的演进诉求，提出 FlexLane 链路高可靠技术构想。该技术基于高速接口多通道架构的现状，打破原有固定组合，引入灵活多通道架构，通过降速运行实时有效的规避任何通道发生的故障，将链路可靠性提升万倍以上（助力 AI 网络互联可靠性超越 5 个 9），保障 AI 训练和推理业务不受影响。FlexLane 技术支持在现有设备上通过软件升级快速部署，或升级硬件实现更优的性能，同时可支持主动降速，在链路轻载和空闲期间动态节能，为智算中心提供灵活、经济、高效的可靠性保障。

本白皮书旨在提出中国移动及产业合作伙伴对以太网链路高可靠 FlexLane 技术的愿景、架构设计和能力要求。希望能够为产业在规划设计智算中心网络、网络互联高可靠相关技术、产品和解决方案时提供参考和指引。

本白皮书由**中国移动通信有限公司研究院**主编，**中国信息通信研究院、清华大学、北京邮电大学、华为技术有限公司、中兴通讯有限公司、上海橙科微电子科技有限公司、新华三技术有限公司、锐捷网络股份有限公司、苏州盛科通信股份有限公司、朗美通通讯技术（深圳）有限公司、武汉光迅科技股份有限公司、思博伦通信科技（北京有限公司）、集益威半导体（上海）有限公司、成都新易盛通信技术股份有限公司、索尔思光电、武汉华工正源光子技术有限公司、上海云脉芯联科技有限公司**联合编撰。

本白皮书不包含我国科技发展战略、方针、政策、计划等敏感信息。不包含涉密项目的背景、研制目标、路线和过程，敏感领域资源、数据，关键技术诀窍、参数和工艺信息。

目　录

1 背景与需求

近年来，人工智能（AI）技术取得了突破性进展，特别是以 ChatGPT、Deepseek 为代表的大语言模型（LLM）的兴起，标志着 AI 进入了一个全新的发展阶段。大模型通常拥有数千亿甚至万亿的参数，中小模型通常也有十亿参数以上，需要海量的算力进行训练和推理。为满足庞大的算力需求，智算中心作为 AI 发展的新型基础设施底座，正加速在全球范围内建设和部署。

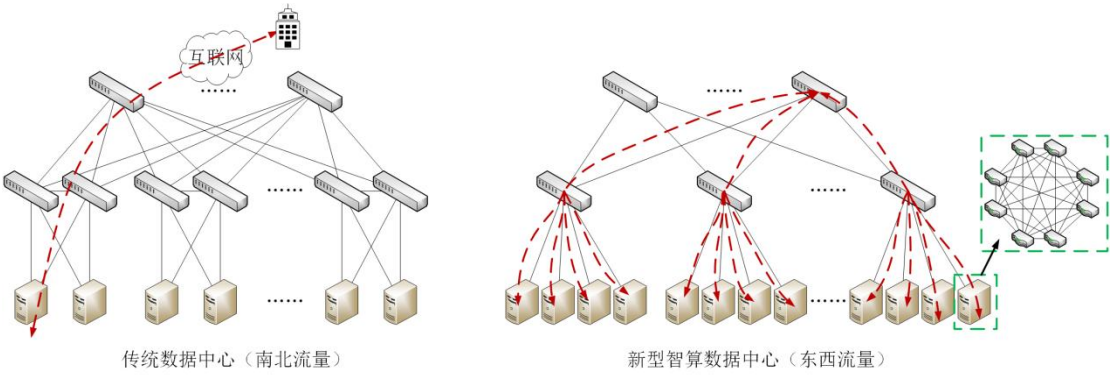

图 1-1 传统数据中心与新型智算中心流量模型对比

传统数据中心主要承载企业级应用，提供云服务，如 Web 应用、数据库、存储等。如图 1-1 所示，这些应用的流量模式以南北向通讯为主，网络的主要任务是保证客户能够及时可靠访问服务器，以及服务器能够快速可靠响应客户请求。用户通过N跳入云，每跳链路的可靠性为R，则业务端到端可靠性为$A = \sum_{i=1}^{N} C_N^i \times R^i \times (1-R)^{N-i} \approx N \times R$（$R = 200FIT$[1]，$N = 3$ 时，$N \times R \approx 6 \times 10^2 FIT$），单个服务器或链路的故障通常只会影响到部分客户端，影响范围相对有限。

新型智算中心主要承载 AI 训练与推理业务，部署大量服务器协同工作，流量模式与传统数据中心不同，东西向流量特征明显。在这种流量模式下，大量服务器共同承载 AI 任务并行计算，对网络的可靠性提出了前所未有的挑战。服务器之间逻辑连接的任何一条物理链路发生故障，都会导致数据同步失败，任务中断，造成大量时间和资源的浪费。如果承载 AI 任务的服务器之间共有N条物理链路，每条链路的可靠性为R，则 AI 训练任务的可靠性为$A = \sum_{i=1}^{N} C_N^i \times R^i \times (1 -$

[1] FIT：Failure in Time of 10^9 hours,在 10^9 小时中发生故障的次数[1]。

$R)^{N-i} \approx N \times R$（$R = 200FIT$，万卡集群无收敛组网$N = 15360$时，$N \times R \approx 3 \times 10^6 FIT$），和传统 DC 业务的可靠性比较，端到端的可靠性下降数千倍以上。根据 Meta LLama 3.1 万卡集群公开的论文[2]，LLama 3.1 在为期 54 天的训练期间共发生 466 次故障中断，其中 GPU、网络互联和主机等故障占比靠前，其中因网络设备和线缆问题造成网络互联故障共 35 次。

光互联链路在带宽、延迟、传输距离等方面具备较大优势，已在智算中心得到广泛部署，如图 1-2 所示[2]。

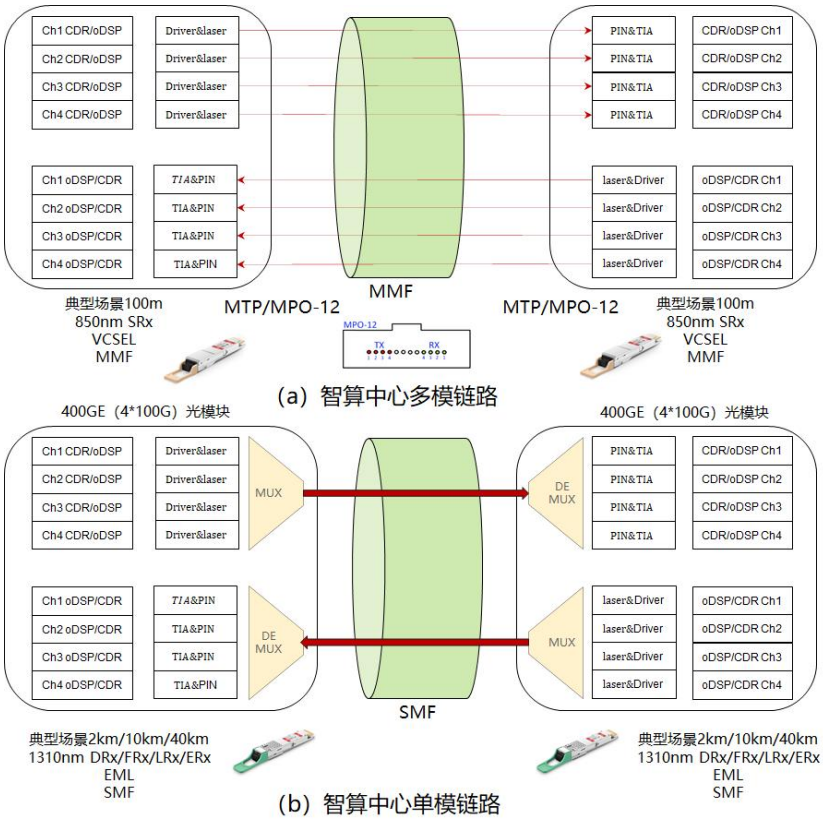

图 1-2 智算中心互联光链路类型

主流高速接口 400G/200G 光模块年失效率超 0.2%，千卡以上集群平均每年发生数十次光模块故障事件。除了器件失效，设备侧或配线架光纤端面脏污也会引发链路闪断[4]，如图 1-3 所示。

[2] 常见多模或单模光模块常为多通道架构，每通道含 CDR（时钟数据恢复，Clock and Data Recovery），DSP（数据信号处理器，Digital Signal Processor）以及激光器等元器件。

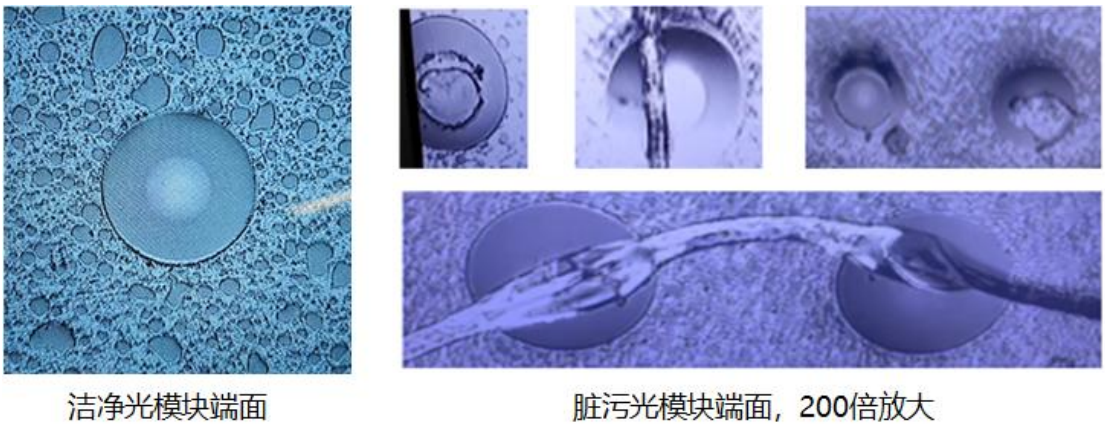

图 1-3 光模块脏污遮挡

链路发生中断或闪断故障会对 AI 训练和推理业务产生诸多影响[5-8]，主要体现在 AI 训练的效率、稳定性和结果准确性，同时也威胁到 AI 推理的可用性、实时性和可靠性。根据业界当前情况，链路故障可能会导致小时级的业务中断。

IEEE802.3 标准以太网[9]面向接口性能最优设计，单一物理通道故障则整条高速链路失效。一个含N个物理通道的标准高速接口故障的概率为：

$$F_{Port} = \sum_{i=1}^{N} C_N^i \times (1 - F_{Lane})^{N-i} \times F_{Lane}^i \approx N \times F_{Lane}。$$

典型的单通道光模块可靠性F_{Lane}约为 100~500FIT[1]，则双通道光模块的标准接口（$N = 2$，$F_{Lane} = 100FIT$）可靠性（1 小时内发生故障的概率）为：

$$F_{Port} \approx N \times F_{Lane} = 2 \times 100 \times 1 \times 10^{-9} = 2 \times 10^{-7}$$

标准接口下的双通道光模块链路在一小时中发生故障的概率为：

$$F_{Link} = \sum_{i=1}^{2} C_2^i \times (1 - F_{Port})^{2-i} \times F_{Port}^i \approx 2 \times F_{Port}$$

$$= 2 \times 2 \times 10^{-7} = 4 \times 10^{-7},$$

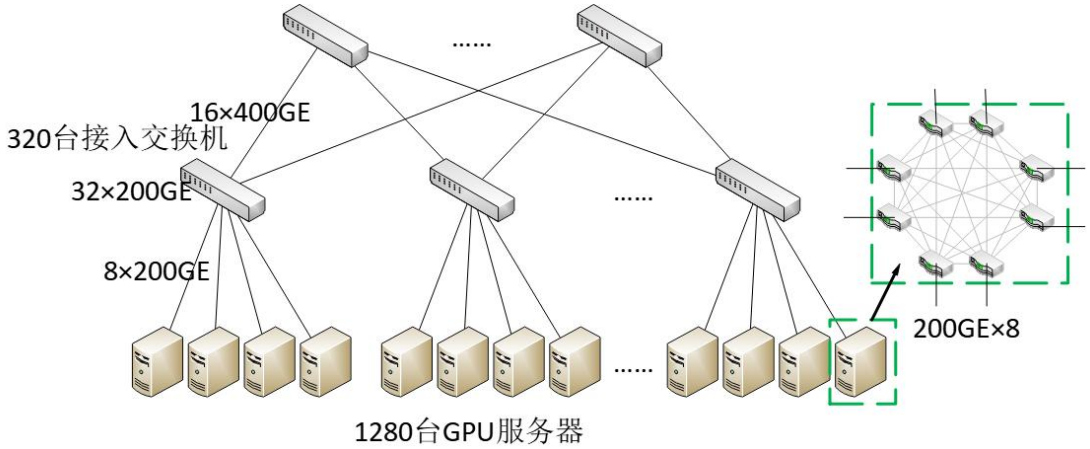

图 1-4 万卡集群示例（10240GPU+15360 链路）

如图 1-4 所示，一个典型的万卡集群无收敛组网，（GPU 总数为 10240，高速互联链路总数 M 为 15360 条），组网中任一链路发生故障会导致网络故障，每小时全网发生故障的概率为：

$$F_{Network} = \sum_{i=1}^{M} C_M^i \times (1 - F_{Link})^{M-i} \times F_{Link}^i \approx M \times F_{Link}$$

$$= 15360 \times 4 \times 10^{-7} = 6.14 \times 10^{-3}$$

根据当前常见大模型披露的训练时间[3]，如表 1-1 所示，在一个万卡集群内，使用标准接口进行大模型训练，过程中发生链路故障的次数约为 2~22 次，无法满足新型智算中心 AI 业务零中断新需求。

表 1-1 使用标准接口进行 AI 大模型训练期间发生链路故障次数

接口类型	M	N_{GPU}	L_{AI_Networ}	$T_{trainning}$[5](hour)	N_{link_fail}[6]
标准接口	15360	10240	约80%（典型值）	334.48（Deepseek-R1）	2.10
				839.80（LLama3.3 70B）	5.16
				3700.00（LLama3.1 405B）	22.71

有多种路径可以实现 AI 业务零中断。就提升光链路可靠性而言，可以在服务器与交换机、交换机与交换机之间广泛部署 LAG 冗余技术，链路可靠性可提升千倍（光模块年失效率 0.4%，光链路年失效率 0.8%，LAG 链路年失效率 0.0016%）。

[3] DeepSeekAI 官方披露是 278.8 万个 H800 小时，LLama3.3 70B 的训练时间是 700 万个 H100 小时，LLama 3.1 405B 是训练了 3084 万个 H100 小时[10]。

[4] $L_{AI\_Network}$：AI 集群网络并行计算线性度。

[5] $T_{trainning}$：万卡 AI 集群网络完成一次大模型训练的时间，$T_{trainning} = \frac{T_{all}}{N_{GPU} \times L_{AI\_Network}}$。

[6] N_{link_fail}：万卡 AI 集群完成一次大模型训练过程中发生链路故障的次数，$N_{link\_fail} = T\_trainning \times F_{network}$

就高速光链路自身而言，单通道失效（器件失效、脏污）占比大，单通道失效阻塞整条链路，资源严重浪费。业界亟需探索新的可靠机制，支持抗单通道或少数通道故障，保障 AI 任务继续运行。

针对上述新型智算中心高可靠承载 AI 业务的诉求，中国移动联合业界合作伙伴提出弹性容错 FlexLane 技术方案，在物理层引入灵活多通道架构，打破原有高速接口与物理通道的固定组合，在单通道或少数通道故障情况下，通过隔离任何故障通道降速工作，可有效提升链路可靠性百万倍以上，确保 AI 任务不因网络互联故障而中断。本白皮书的发布有望推动 FlexLane 技术的产业共识、技术成熟与商用落地，支撑智算中心的 AI 训练和推理业务稳定运行与发展。

2 FlexLane 技术架构

2.1 技术目标

FlexLane 物理层方案更便于实现高可靠、低时延、低开销的保障能力，预期可避免网络互联故障，保障 AI 任务零中断，满足智算中心场景对网络的要求。

2.2 设计原则

2.2.1 兼容性原则

FlexLane 技术可以在网络的不同层级位置实现。在物理层 PHY 单元实现时，要求兼容已有标准（例如 IEEE802.3），不影响标准已规范的功能与协议。在上层软件实现时，要求兼容通用的网络协议栈，并保持与现有应用的兼容性。FlexLane 技术与上层可靠性方案，例如 RDMA 重传、LAG 等可同时部署。

2.2.2 一致性原则

面向标准规范，例如 IEEE802.3 规范的高速以太网 100GE/200GE/400GE/800GE/1.6TE 接口，提供一套 FlexLane 技术架构和协议。同一层次方案，要求协议一致，满足互联互通要求。

2.3 技术架构

本白皮书提出的高可靠方案部署层级架构如图 2-1(a) 所示，以 200GE/400GE/800GE 为例说明，技术架构主要包括三个关键子系统：检测功能、切换机制和交互协议，如图 2-1(b)所示：

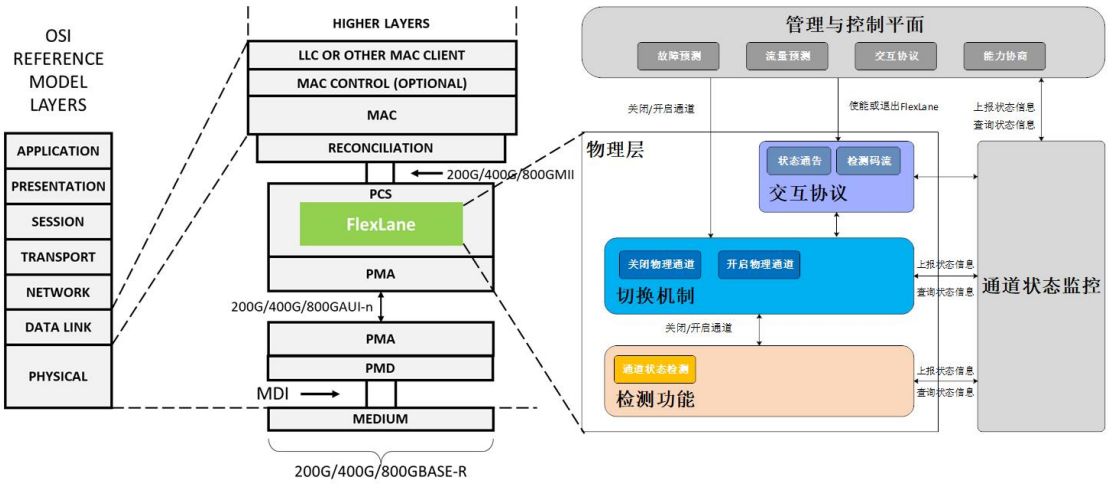

图 2-1 FlexLane 技术架构与部署层级

- 交互协议：链路两端通过协议报文向对端通告故障隔离、故障恢复等操作。
- 切换机制：管理物理通道的状态（开启/关闭）。当检测到故障时，支持隔离故障通道；当检测到故障通道恢复正常后，支持将故障通道恢复为正常工作通道。支持主动开启或关闭部分通道实施故障预防策略（例如上层应用提前诊断出某通道即将发生故障），或动态节能。
- 检测功能：实时检测各通道状态。支持被动查询或主动上报物理通道的状态，含发光功率、收光功率、温度、电流、电压等信息。

FlexLane 的协商协议、切换机制以及检测功能都可以与更上层的管控系统进行交互，从而对通道的状态进行监控，如查询通道当前信号质量、通道当前状态（正常工作/故障/恢复中）以及当前流量特征等通道管控操作。应用接与控制平台也可以主动对通道进行管理与控制，如下发指令关闭/开启某通道。

3 FlexLane 关键技术

FlexLane 的整体流程包含故障通道的检测、故障通道隔离、故障通道恢复以及主动开启或关闭通道，如图 3-1 所示。

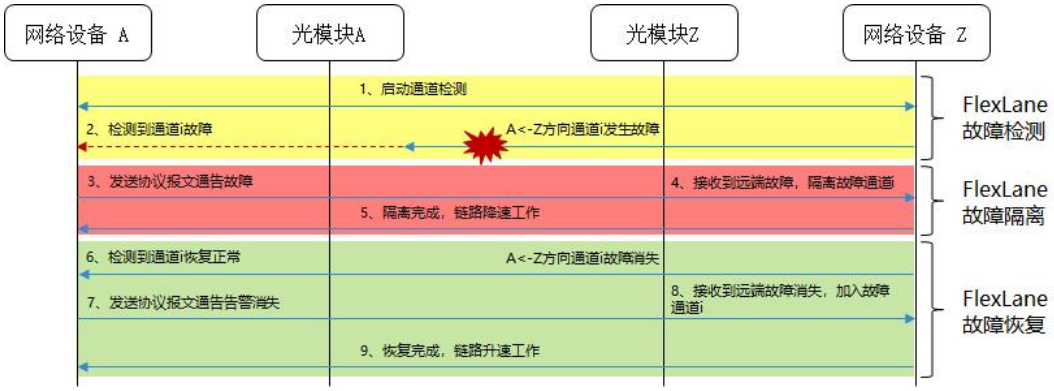

图 3-1 FlexLane 整体技术流程

- 故障检测：本端支持通道粒度的告警检测。高速接口的任一通道发生故障时，立即触发故障隔离流程。
- 故障隔离：本端向远端发送故障信息协议信令，通知远端隔离发送侧对应故障通道。同时启动本地故障通道隔离流程，停止从故障通道接收信息。远端收到故障信息协议信令，停止往故障通道发送信息。故障隔离完成，接口降速运行。
- 故障恢复：通道故障消失后，接收侧向远端发送故障消失的协议信令。本地和远端启动恢复流程，被隔离通道重新加入链路工作。

此外，还需支持主动降速/升速模式，由管理或控制平面触发 FlexLane 降速，主要用于如下场景：

- 应用层检测到某通道的信号正在劣化，提前下发降速指令规避故障发生；
- 应用层预测到未来高速链路流量将会轻载甚至空闲，手工下发降速指令，关闭部分通道耗能元器件动态节能。

3.1 故障隔离

针对现网情况，可考虑 FlexLane 的灵活部署策略，如图 3-2 所示：近期通过软件升级支持 FlexLane，可快速部署；面向未来，选择在高速接口硬件实现，可

获得最佳性能。

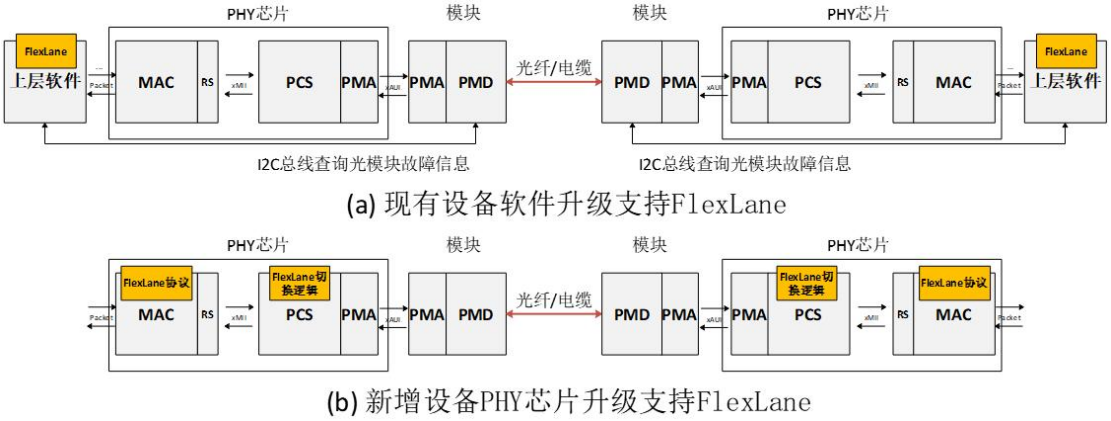

图 3-2 FlexLane 灵活部署策略

3.1.1 软件升级

FlexLane 软件方案升级网络设备和光模块软件，支持通道状态查询和上报，在不更换硬件的情况下实现故障通道隔离。故障检测和通道隔离由上层软件发起，如图 3-3(a)所示，以四通道 400GE 高速接口为例，

- ✓ T1 时刻，Z 端发送侧光模块发生闪断或中断故障；
- ✓ T2 时刻，A 端上层软件检测到链路故障，通过接口查询 PHY 芯片或者光模块后获取故障通道信息；
- ✓ T3 时刻，A 端上层软件通过软件协议通告故障信息；
- ✓ T4 时刻，Z 端上层软件根据故障信息，发送握手信息约定隔离完成的边界，并在握手信息发送完成后将发送侧接口重配置降速（例如降速为 200GE 运行），如图 3-5(b)所示；
- ✓ T5 时刻，A 端上层软件在握手信息接收完成后将接收侧接口重配置降速。

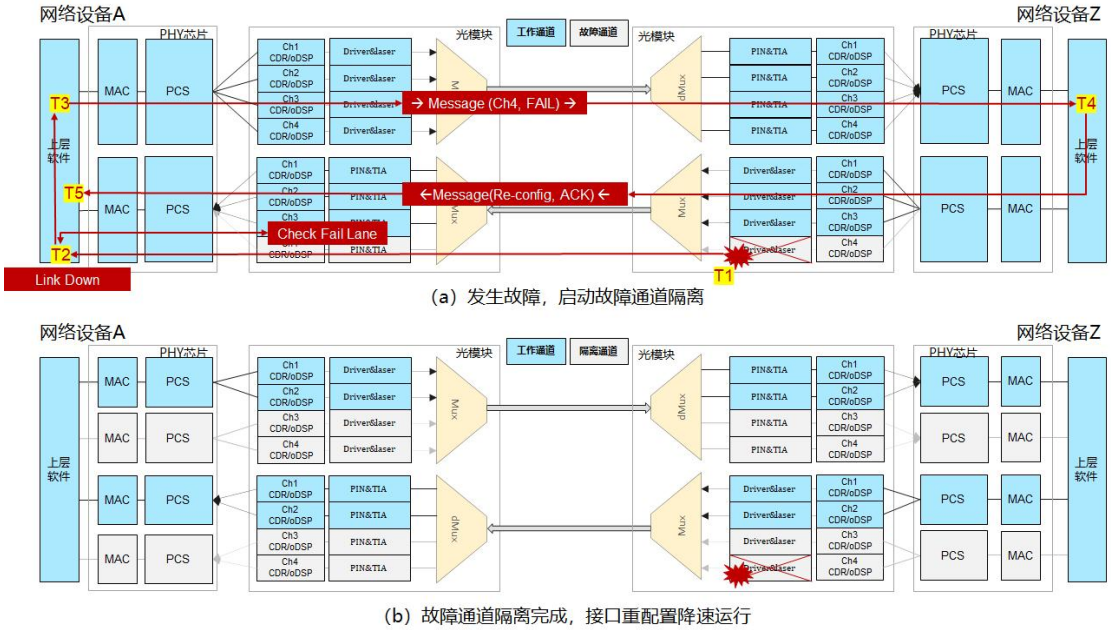

图 3-3 通道隔离软件方案 400GE 降速为 200GE 流程示意

当网络设备上层软件检测到链路恢复，支持对端口进行重配置升速，恢复带宽以获得更佳的计算效率。

3.1.2 硬件演进

FlexLane 硬件演进方案升级 MAC/PHY 接口，新增物理层故障检测能力。硬件方案的故障检测、通道隔离和通道恢复流程由物理层 FlexLane 协议发起。

（1）故障检测

通道故障类型可分为单通道故障（单向单通道故障和双向单通道故障）、多通道复杂故障，其中单向单通道故障发生的频率最高，如图 3-4 所示。

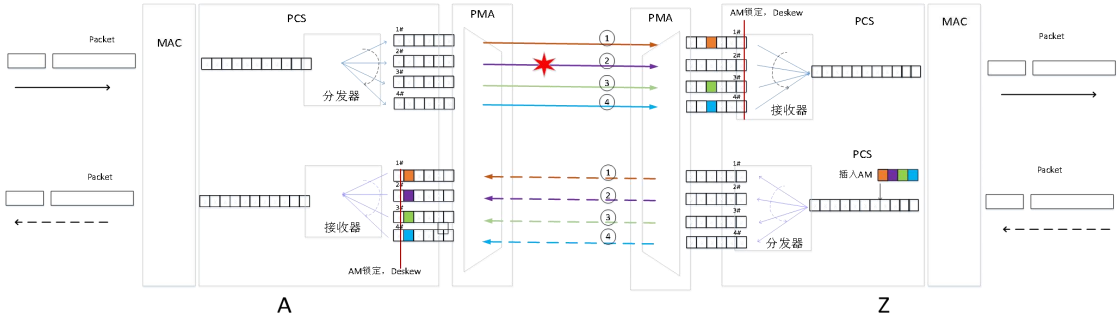

图 3-4 单向单通道故障（典型故障）

针对链路信号丢失故障（SF，Signal Failure），采用基于通道 AM 失锁检测方

案（参考 802.3 CL 119.2.6.3，2022），如图 3-5 所示，当检测到某个通道连续N个 AM 周期丢失锁定（N 缺省 5），判断该通道失效，进而引发链路 SF。

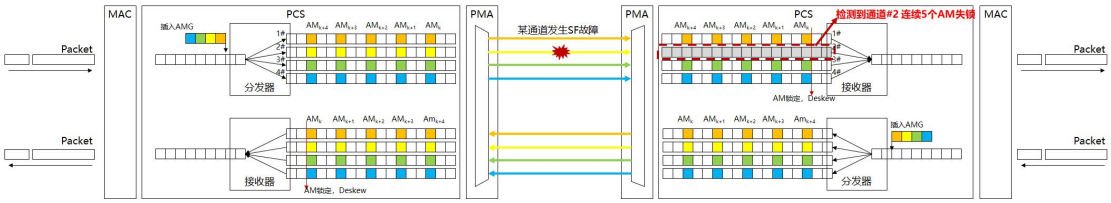

图 3-5 SF 故障检测机制

针对链路信号质量差导致的故障（SD，Signal Degrade），采用符号错误率（SER，Symbol Error Rate）统计方案（参考 802.3 CL 119.2.5.3，2022）；或如图 3-6 所示，统计各通道的 SER（缺省窗口 8192 个 FEC CW），当一个通道 SER 超过阈值时（缺省 5560 个 Symbol）,认为该通道发生 SD 故障。

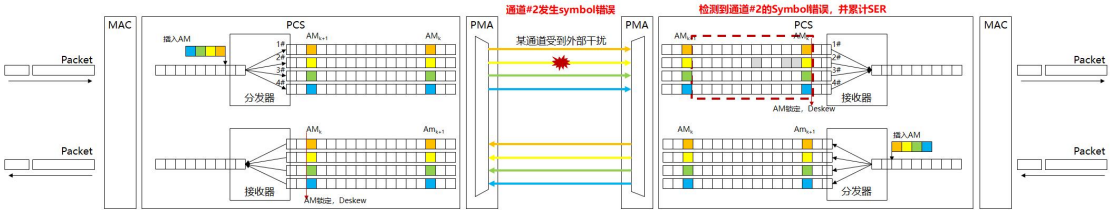

图 3-6 SD 故障检测机制

针对链路信号质量监测，可以基于光模块状态参数，例如温度、电压、电流、接收光功率、发送光功率，结合通道 SER 做统一预测和决策，根据决策结果可以主动发起对存在风险的通道提前隔离。

（2）故障隔离

FlexLane 支持在检测到故障发生后对故障通道进行隔离，避免故障导致 AI 业务中断。故障隔离机制如图 3-7 所示，以四通道 400GE 高速接口为例，

- ✓ T1 时刻，Z 端发送侧的某一光模块发生闪断或中断故障；
- ✓ T2 时刻，A 端检测到 SF 或 SD 故障，立即隔离故障通道，并停止在所有通道上接收业务数据流；
- ✓ T3 时刻，A 端发送协议报文通告故障信息。
- ✓ T4 时刻，Z 端收到故障信息并对隔离故障通道，停止在所有通道上发送业务数据流。
- ✓ T5 时刻，Z 端发送握手信息约定故障隔离完成后业务恢复的切换边界，并在握手信息发送完成后重新在正常工作的通道上发送业务数据流。

✓ T6 时刻，A 端在握手信息接收完毕后重新在正常的通道上接收业务数据流。

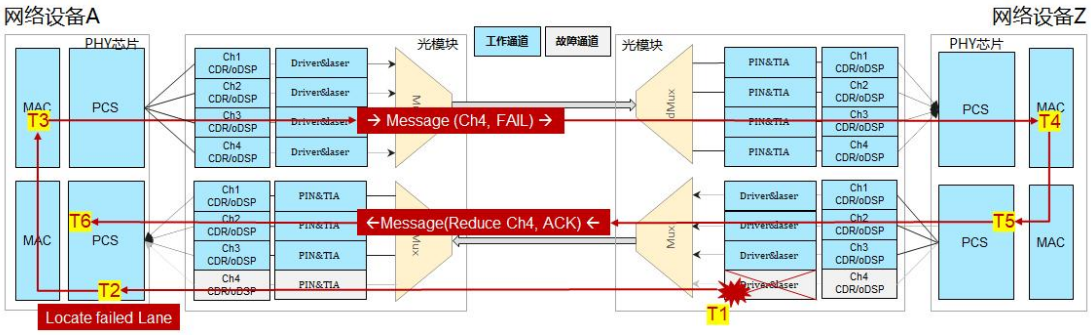

图 3-7 通道隔离硬件方案流程示意图

FlexLane 支持在通道发生劣化但未故障时进行主动降速，避免故障发生。当接收到上层应用（包括管理面和控制面）主动隔离某通道降速运行的命令，实施流程如图 3-8 所示，以四通道 400GE 高速接口为例，

✓ T1 时刻，当 A 端的上层软件检测到某通道未来存在故障风险，决定对该通道实施主动隔离，执行降速操作，向 Z 端发送协议报文，通告关闭对应通道；

✓ T2 时刻，Z 端的上层软件收到通告报文，向 A 端发送握手信息约定主动关闭对应通道的切换边界；

✓ T3 时刻，Z 端发送握手信息完成后，关闭对应通道，停止在该通道上分发数据流；

✓ T4 时刻，A 端接收握手信息完成后，关闭对应通道，停止在该通道上接收数据流。

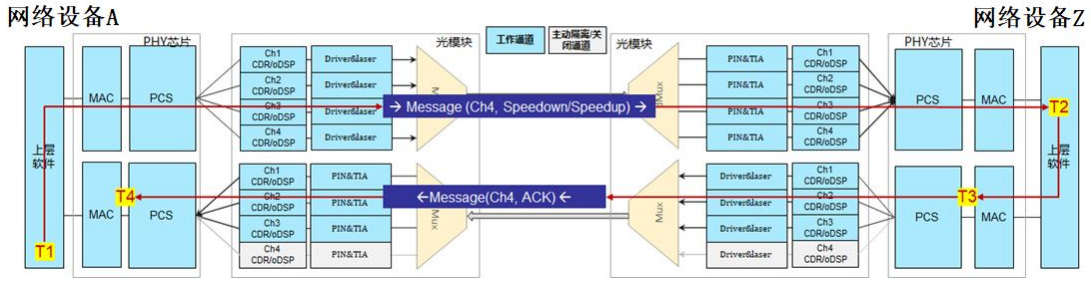

图 3-8 主动升速/降速方案流程示意图

主动降速机制同时支持链路轻载或空载时，主动关闭部分通道，实施动态节能；当流量即将满载或重载时，再全速运行。

（3）故障恢复

FlexLane 支持将故障消失的通道重新加入高速接口，恢复带宽，获取更高的运行效率。通道恢复的关键技术在于如何检测隔离后通道的状态，以及如何保证

无损增加通道，以图 3-9 为例说明，

- ✓ T1 时刻，Z 端在完成通道隔离后，持续向 A 端的故障通道上发送协议报文（例如发送 IDLE 码块）；
- ✓ T2 时刻，故障消失，A 端可以接收到正常的协议报文（IDLE 码块），探测到通道的故障已消失；
- ✓ T3 时刻，A 端发送协议报文通告故障消失信息；
- ✓ T4 时刻，Z 端收到故障通道恢复的协议报文；
- ✓ T5 时刻，Z 端发送握手信息约定故障通道恢复成工作状态的切换边界（例如 AM），并在握手信息发送完成后重新在正常工作的通道上发送业务数据流。
- ✓ T6 时刻，A 端在握手信息接收完成后重新在正常的通道上接收业务数据流。

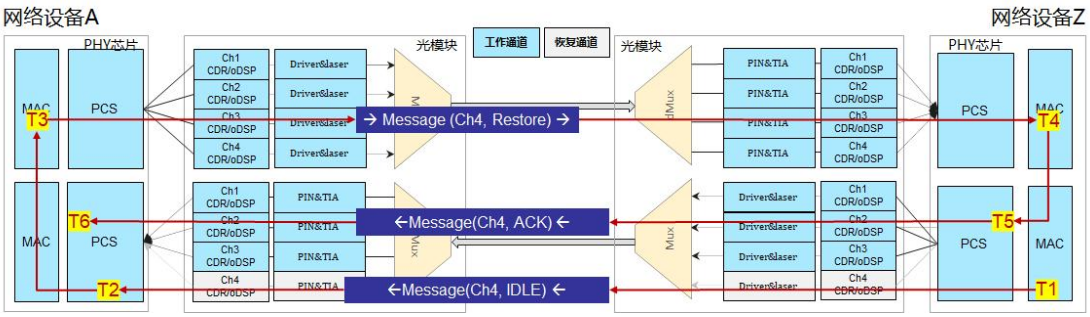

图 3-9 通道恢复硬件方案流程示意图

3.1.3 技术效果

FlexLane 接口支持隔离任意故障通道降速运行，只有当接口中的所有通道都故障时，接口才会失效，一个含N个通道的 FlexLane 接口故障失效的概率为：

$$F_{Port} = \sum_{i=N}^{N} C_N^i \times (1 - F_{Lane})^{N-i} \times F_{Lane}^i \approx F_{Lane}^N$$

典型的单通道光模块可靠性F_{Lane}约为 100~500FIT，则双通道光模块的 FlexLane 接口（$N = 2$，$F_{Lane} = 100FIT$）可靠性（1 小时内发生故障的概率）为：

$$F_{Port} \approx F_{Lane}^N = {F_{Lane}}^2 = (100\times 1 \times 10^{-9})^2 = 1 \times 10^{-14}$$

FlexLane 接口下的双通道光模块链路在一小时中发生故障的概率为：

$$F_{Link} = \sum_{i=1}^{2} C_2^i \times (1 - F_{Port})^{2-i} \times F_{Port}^i \approx 2 \times F_{Port}$$

$$= 2 \times 1 \times 10^{-14} = 2 \times 10^{-14},$$

如图 2-2 所示，一个典型的万卡集群无收敛组网，（GPU 总数为 10240，高速互联链路总数 M 为 15360 条），使用 FlexLane 接口，每小时全网发生故障的概率为：

$$F_{Network} = \sum_{i=1}^{M} C_M^i \times (1 - F_{Link})^{M-i} \times F_{Link}^i \approx M \times F_{Link}$$

$$= 15360 \times 2 \times 10^{-14} = 3.07 \times 10^{-10}$$

使用 FlexLane 或标准以太接口的集群网络，在一小时内发生链路故障的概率如表 3-1 所示，使用 FlexLane 接口的集群网络在一小内发生链路故障的概率比使用标准接口的情况下低 7 个数量级。

表 3-1 标准接口与 FlexLane 接口链路可靠性

接口类型	N_{link}	F_{Lane}	F_{Port}	F_{Link}	$F_{Network}$
标准接口	15360	100FIT	2.00E-07	4.00E-07	6.14E-03
FlexLane接口			1.00E-14	2.00E-14	3.07E-10

根据当前常见大模型披露的训练时间，如表 3-2 所示，在一个万卡集群内，使用 FlexLane 接口进行大模型训练，过程中发生链路故障的次数比使用标准小 7 个数量级，AI 网络光互联部分的可靠性可达 6 个 9。

表 3-2 使用标准接口和 FlexLane 接口进行 AI 大模型训练期间发生链路故障次数对比

接口类型	M	N_{GPU}	L_{AI_Networ}	$T_{trainning}$	N_{link_fail}
标准接口	15360	10240	约80%（典型值）	334.48（Deepseek）	2.10
				839.80（LLama3.3 70B）	5.16
				3700.00（LLama3.1 405B）	22.71
FlexLane 接口				334.48（Deepseek）	1.03E-07
				839.80（LLama3.3 70B）	2.58E-07
				3700.00（LLama3.1 405B）	1.14E-06

3.2 故障预防

FlexLane 支持在故障发生前关闭劣化通道，避免 AI 任务因故障发生中断，实

现无损的数据传输。当某通道信号逐渐劣化（如在一个时间窗内，错误符号率 SER 超过特定阈值）但未触发 SD 故障时，FlexLane 可以主动上报，并根据控制器或网管平面决策实施预防策略（例如重启、重训练等），避免 SD 故障发生。

3.3 动态节能

FlexLane 的主动降速升速机制支持根据信道质量和流量变化，关闭或开启接口中的通道。结合目前业界的商用部署情况，高速接口(MAC/PHY 和 SerDes)普遍占交换机主芯片能耗约 50%，当链路处于低流量场景时，可通过关闭部分通道的耗能元件（分布于 MAC/PHY、SerDes 和光模块）降低能耗。

新型智算中心场景下，AI 大模型训练过程中的流量模型具有方波性，如图 3-10 所示，某 GPT-3 组网，GPU 之间网络利用率约 5%，交换机之前网络利用率仅 1%。AI 集群网络在等待计算期间产生网络互联“空跑”能耗。

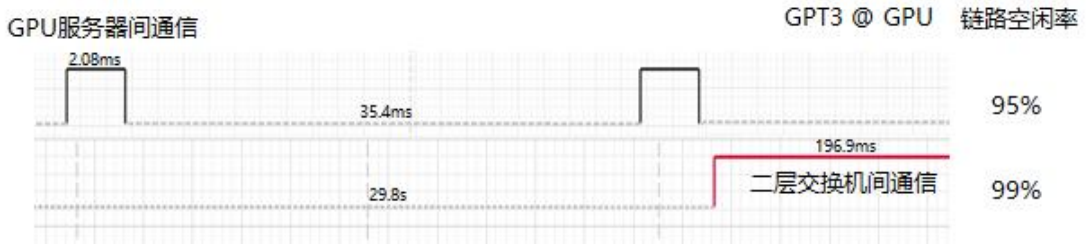

图 3-10 AI 大模型训练网络流量特征示意

以支持四通道的 FlexLane 接口为例，PP 并行，在 GPU 服务器计算期间，互联链路空闲，关闭三个通道保留一通道运行，接口能耗理论上[7]降低$\frac{35.4}{2.08+35.4}\times\frac{3}{4}=70.9\%$，主芯片能耗理论上可降低$50\%\times70.9\%=35.5\%$；DP 并行，在二层网络设备互联空闲期间，每链路关闭三个通道，接口能耗理论上降低$\frac{29800}{29800+196.9}\times\frac{3}{4}=74.5\%$，主芯片能耗理论上可降低$50\%\times74.5\%=37.3\%$。

[7] 实际节省的能耗比例与器件能力及实施策略有关。

4 应用场景

FlexLane 技术可广泛应用于移动承载、园区、智算中心以及中心间互联各种组网场景。

对于移动承载等场景，连通性对网络稳定运行影响大，实施 FlexLane 技术，可彻底规避由于通道相关的器件引发的连通性故障， 一条含N个通道的链路，使用 FlexLane 接口，可靠性由$N \times F_{Lane}$提升至F_{Lane}^{N}。

对于智算中心或智算中心间互联等场景，带宽损失对计算任务影响较大，考虑链路带宽损失，一条含N个通道的高速链路仅支持降一通道的 FlexLane 策略（带宽仅损失 1/N），端口的可靠性为，

$$F_{Port} = \sum_{N}^{i=2} \mathrm{C}_{N}^{i} \times (1 - F_{Lane})^{N-i} \times F_{Lane}^{i} \approx C_{N}^{2} \times F_{Lane}^{2}$$

可靠性由$N \times F_{Lane}$提升至$\frac{N\times（N-1）}{2} \times F_{Lane}^{2}$

4.1 智算中心

智算中心内服务器与网络设备、网络设备与网络设备高速互联，重点承载 AI 推理与训练任务，对延迟和带宽要求高。AI 任务普遍需多台服务器并行计算，计算期间需频繁交换大量梯度数据和模型参数，网络闪断或中断会影响计算效率。

服务器与网络设备、网络设备与网络设备之间普遍部署短距高速光模块互联（N通道），为了尽量不损失带宽，每链路只支持降一条通道策略（带宽降低 $1/N$），如图 4-1 所示，以典型 400G 光模块（4 条 100G 通道）100m 多模互联链路为例，一个方向发生单通道故障，该方向降速为 300G，另外一个方向仍然维持 400G 运行。当$N = 4$，$F_{Lane} = 100FIT$，1 小时内发生故障的概率由$N \times F_{Lane} = 4 \times 100 \times 10^{-9} = 4 \times 10^{-7}$ 降 低 至 $C_{N}^{2} \times F_{Lane}^{2} = 6 \times （100 \times 10^{-9}）^{2} = 6 \times 10^{-14}$。

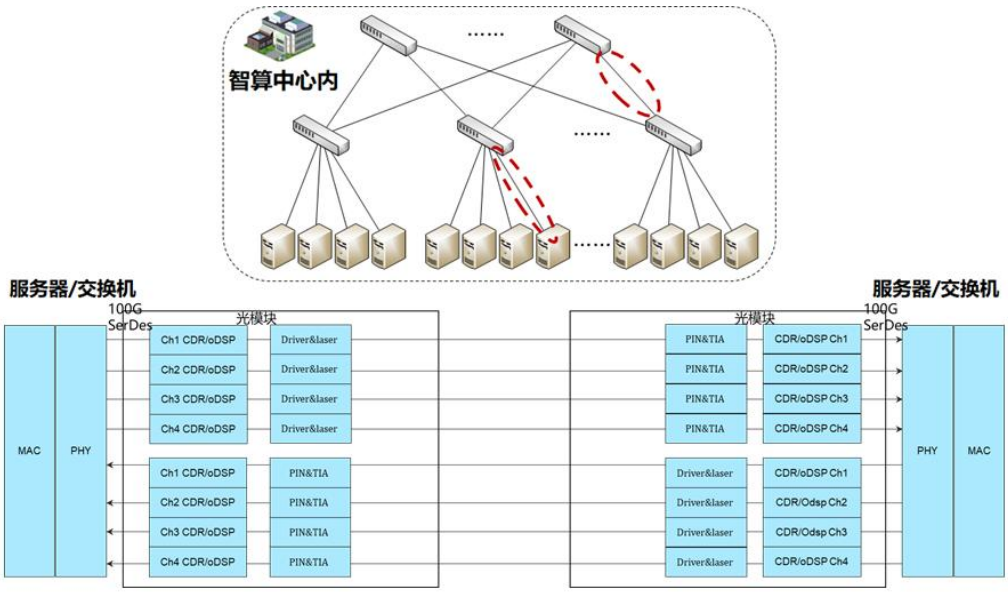

图 4-1 智算中心内部互联链路

4.2 智算中心互联

智算中心间出口网络设备之间部署高速直检光模块互联（N通道）链路，如图 4-2 所示。支持 FlexLane 技术后，互联链路的任一通道的关键器件故障，都不影响连通性，如果只支持降一通道，则对带宽影响也较小；如果支持降到一通道运行，则可靠性将大幅提升：以典型 400GE（4 条 100G 通道）10km 单模互联链路为例，支持三条通道故障隔离后，1 小时内发生故障的概率由$N \times F_{Lane} = 4 \times 100 \times 10^{-9} = 4 \times 10^{-7}$降低至$F_{Lane}^{4} = (100 \times 10^{-9})^{4} = 1 \times 10^{-28}$。如此，部署 10^{12} 条 400GE 链路的超大网络，宇宙年内不会发生因通道器件（LD、PD、TIA 和 Driver 等）故障而丢失连通性。

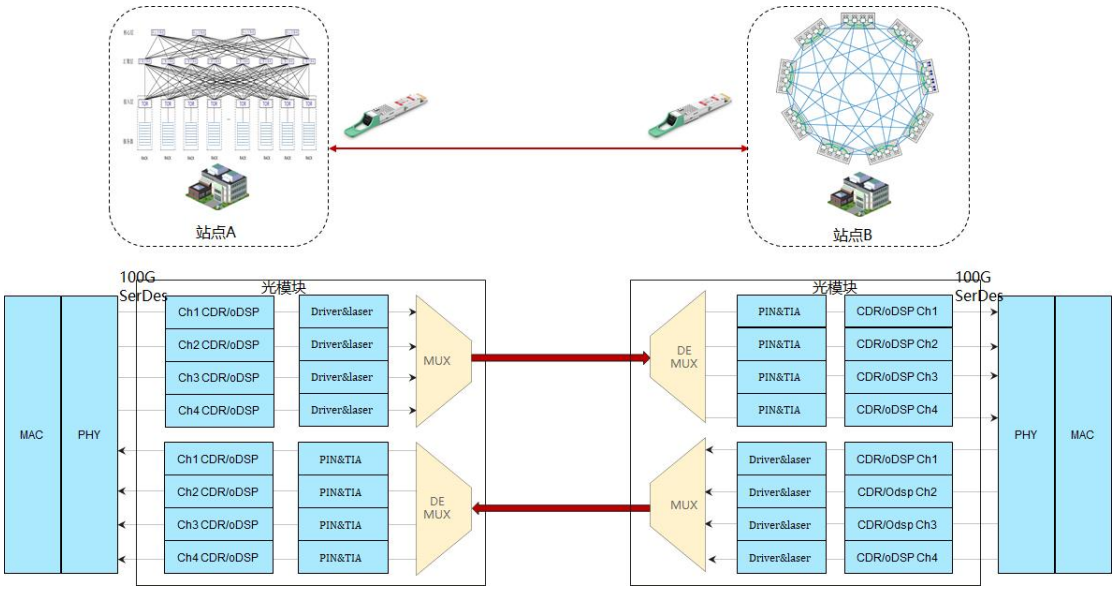

图 4-2 智算中心间高速直检互联链路

智算中心间出口网络设备之间部署高速相干光模块互联（主机侧N通道，线路侧单通道单纤）链路，如图 4-3 所示的 400G ZR+示例，线路侧单通道单纤架构无短距可靠性降N倍问题（$F_{Port} = F_{Port}$），其次相干链路投资大，相干光模块采用高品质器件以及高等级封装，关键器件失效概率和灰尘遮挡概率小；主机侧 4 通道，相比高速直检链路，主机侧电接口及接插件故障在相干链路故障总占比较大。FlexLane 技术可提升链路在主机侧接口的可靠性，任一 SerDes 故障（例如接插件异常），FlexLane 可隔离故障并维持链路继续运行。

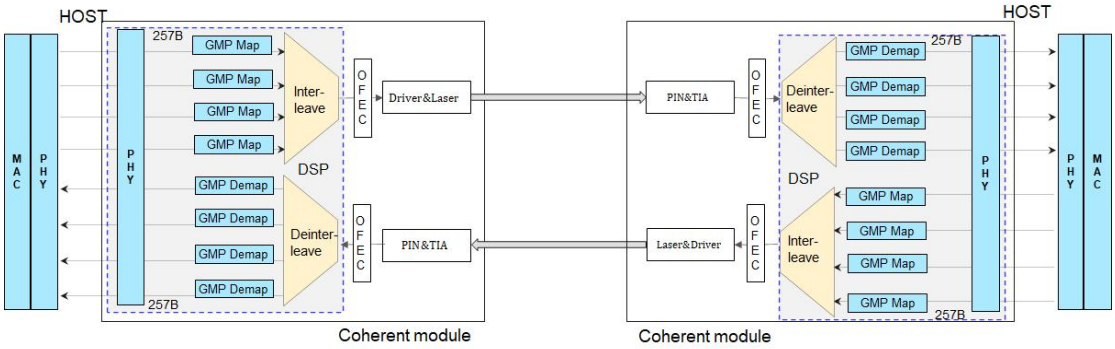

图 4-3 智算中心间高速相干互联链路

5 总结与展望

随着 AI 大模型兴起，智算中心成为全球算力基础设施建设焦点。AI 业务对网络链路可靠性要求极高，网络互联故障将导致任务中断。本白皮书提出的 FlexLane 技术，引入灵活多通道架构，将链路可靠性提升万倍以上（助力 AI 网络互联可靠性超越 5 个 9），确保 AI 任务不因网络互联故障而中断，大幅度提升 AI 基础设施可靠性。

FlexLane 聚焦现有链路挖潜，无需更换高品质光器件，具备低成本优势，支持现有设备软件升级部署，或新设备硬件集成，为互联链路提供灵活、经济和高效的可靠性保障，预计在智算中心内部及间互联场景将获得广泛部署。

FlexLane 技术与链路级重传（LLR）技术结合，可实现高速互联故障无损；链路降速信息实时上报至管理或控制平面，可实现全网调优。展望未来，FlexLane 将持续演进，引导未来高速接口产业走向接口性能最优与可靠性并重，为多通道高速接口扫清障碍，助力智算中心网络互联迈向更大规模领域。

FlexLane 主动降速机制，根据通道的信号质量主动隔离有风险通道，提前规避故障的发生；可以根据流量变化，动态关闭部分通道，在网络轻载或空闲时降低能耗，节能减排。

FlexLane 是极佳的提升智算中心可靠性的低成本解决方案，为未来 1.6TE 及更大带宽应用保驾护航，同时兼顾动态节能特性，有望与业界尽快达成共识，广泛部署。

缩略语列表

缩略语	英文全名	中文解释
AI	Artificial Intelligence	人工智能
AM	Alignment Marker	对齐操作码块
CDR	Clock and Data Recovery	时钟数据恢复
DC	Data Center	数据中心
DSP	Digital Signal Processor	数据信号处理器
FEC	Forward Error Correction	前向纠错码
FIT	Failure in time of 10^9 hours	十亿小时发生错误次数
GPU	Graphic Processing Unit	图形处理器
HBM	High Bandwidth Memory	高带宽内存
LAG	Link Aggregation Group	链路聚合
LD	Laser diode	激光二极管
LLM	Large Language Model	大语言模型
LLR	Link Level Retransmission	链路级重传
MAC	Media Access Control Layer	介质访问控制层
PCS	Physical Coding Sublayer	物理编码子层
PD	Photodiode	光电二极管
PHY	Physical	物理层
PMA	Physical Medium Attachment	物理媒介适配层
RDMA	Remote direct memory access	远程直接内存访问
SD	Signal Degrade	信号劣化故障
SER	Symbol Error Rate	符号错误率
SerDes	Serializer/Deserializer	串行器/解串器
SF	Signal Failure	信号丢失故障
TIA	Trans-impedance amplifier	跨阻放大器

参考文献

[1] Texas Instruments. 了解符合 IEC 62380 和 SN 29500 的功能安全时基故障基本 故障率估算.(2020). https://www.ti.com.cn/cn/lit/wp/zhcaaa7a/zhcaaa7a.pdf?ts=1743666213645&ref_url=https%253A%252F%252Fwww.google.com%252F

[2] AI Meta. The Llama 3 Herd of Models. (2024). https://ai.meta.com/research/publications/the-llama-3-herd-of-models/

[3] 中国移动. 中国移动 NICC 新型智算中心技术体系白皮书. (2023)

[4] 腾讯云. 光纤端面验证. (2020). https://cloud.tencent.com/developer/news/694463

[5] Huawei. Atlas 800T A2 训练服务器 维护与服务指南 15. (2024)

[6] OPT: Open Pre-trained Transformer Language Models. (2022). https://arxiv.org/abs/2205.01068

[7] GPT-4「炼丹」指南：MoE、参数量、训练成本和推理的秘密. (2023)

[8] Ali cloud. 大模型训练稳定性思考和实践. (2024)

[9] IEEE. IEEE Standard for Ethernet. (2022). https://www.ieee802.org/3/

[10] AI 大模型学习. 20 条关于 DeepSeek 的 FAQ 解释 DeepSeek 发布了什么样的模型？为什么大家如此关注这些发布的模型？他们真的绕过 CUDA 限制，打破了 Nvidia 的护城河了吗？. (2025). https://www.datalearner.com/blog/1051738420483556

[11] 中国移动. 全调度以太网技术架构（GSE）白皮书. (2023)