

# 易飞扬对工具智能和国家算力网的组网技术白皮书

发布单位：易飞扬（GIGALIGHT）

发布日期：2026 年 07 月

文档版本：V0

## 目录

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| 引言 .....                                                                | 2  |
| 一、 采用 NRZ 调制或半 DSP 调制打造低时延和低功耗的算力数据中心内部网络 .....                         | 2  |
| 1.1 NRZ 调制：成熟可靠，极致低时延 .....                                             | 2  |
| 1.2 “半 DSP” 调制方案：LPO 与 Hybrid 技术 .....                                  | 3  |
| 二、 采用 O-BAND 非相干 DWDM 作为 DCI 相干通信的有力补充，实现低时延低成本中短距 DCI .....            | 6  |
| 三、 1.6T AI&DC 互连的三种灵活解决方案 .....                                         | 8  |
| 3.1 标准 1.6T 可插拔光模块（Pluggable Transceiver） .....                         | 8  |
| 3.2 外置光源的 1.6T 可插拔光模块（External Light Source Pluggable，ELS-Pluggable） .. | 9  |
| 3.3 外置光源的 1.6T（16×100G）NPO 解决方案（Near-Package Optics，NPO） .....          | 10 |
| 3.4 三种方案对照比较 .....                                                      | 11 |
| 四、 采用 AWGR+EDFA/AWGR+SOA 实现关键节点枢纽的全光低时延交换 .....                         | 12 |
| 4.1 AWGR（阵列波导光栅路由器） .....                                               | 12 |
| 4.2 AI 算力集群 AWGR+EDFA 拓扑结构 .....                                        | 12 |
| 4.3 AI 算力集群 AWGR+SOA 拓扑结构 .....                                         | 13 |
| 4.4 关键节点选型建议： .....                                                     | 14 |
| 结束语 .....                                                               | 15 |

## 引言

算力是工具而非智能。我们应当重新定义一种 “工具智能”。亚里士多德《工具论》并非面向日常用具，而是定义范畴与逻辑。范畴和逻辑是思考之上的思考，依托工具思维，才能走向科学与技术创新的正确路径。

全国一体化国家算力网需要一套具备长期兼容、平滑迭代的底层技术体系。若算力基础设施不具备原生先进性，会在技术迭代中快速淘汰，形成大量重复建设与资源浪费。传统臃肿、高消耗的组网架构无法支撑长期算力调度与工具智能释放，因此算力网必须打造轻量化、低时延、低成本的原生架构，相比海外网络体系形成天然竞争优势。

算力网络与人类大脑运行逻辑相似：架构臃肿、处理开销大则算力调度 “停滞”，只有轻量化全光互联架构，才能持续释放算力的工具智能价值。

## 一、采用 NRZ 调制或半 DSP 调制打造低时延和低功耗的算力数据中心内部网络

在算力数据中心内部网络中，采用 NRZ（Non-Return-to-Zero，不归零）调制以及 “半 DSP” /线性驱动方案（如 LPO、Hybrid）已成为兼顾低时延与低功耗的重要技术路径。两类方案分别面向不同的速率等级与传输距离场景，核心目标是在满足高速互联需求的同时，缓解传统全 DSP 方案功耗高、时延大的问题。

### 1.1 NRZ 调制：成熟可靠，极致低时延

NRZ 是一种传统的二进制调制方式，每个符号携带 1 比特信息。在短距离互联中，它凭借简单的

物理层特性成为低时延首选。

**适用场景：**主要应用于单通道 $\leq 200\text{Gbps}$  速率的互联场景。随着技术发展，部分高性能 SerDes 也在探索基于 NRZ 原理的优化方案用于更高速率，但主流高速率（400G/800G）已转向 PAM4。

**核心优势：**

- 1. 极低时延：**无需复杂的数字信号处理（DSP）进行均衡和纠错，信号从电到光的转换几乎无额外处理延迟。
- 2. 低功耗：**电路结构简单，没有 DSP 芯片的高功耗负担，单模块功耗通常控制在较低水平（如几瓦特以内）。
- 3. 高信噪比容忍度：**在光纤中，信道损耗较小，NRZ 的二电平信号具有较好的抗干扰能力。

**技术局限：**频谱效率较低，难以支撑单波长 50Gbps 以上的高速传输，因此在 400G/800G 及以上速率的主流数据中心和 AI 算力架构互联中逐渐被 PAM4 取代。

易飞扬的 200G QSFP-DD S8/PSM8 平行光模块是这个领域的里程碑。

## 1.2 “半 DSP” 调制方案：LPO 与 Hybrid 技术

所谓的“半 DSP”并非标准术语，通常指代弱 DSP+ LPO（Linear Pluggable Optics，线性可插拔光学）或 Hybrid/LRO（Linear Receive Only）等去 DSP 化或弱 DSP 化的技术方案。这些方案旨在保留一定信号完整性的同时，大幅降低传统全 DSP 方案的功耗和时延。

### LPO(Linear Pluggable Optics)

LPO 方案完全移除了光模块中的 DSP 芯片，仅保留线性驱动器（Driver）和跨阻放大器（TIA），依赖交换机侧 ASIC 芯片的强大均衡能力来处理信号。

### Hybrid（Half Retimed & half LRO）/ LRO (Linear Receive Only)

Hybrid 方案通常指在发射端或接收端保留部分线性驱动功能，或在特定方向上使用简化的 DSP 逻辑，以平衡性能与功耗。

### 1. 特点：

- 功耗与性能的平衡：相比 LPO，Hybrid 方案可能在长距离（如 100m-2km）上提供更好的信号完整性，同时仍比全 DSP 方案节能 20%-30%。
- 时延优化：虽然保留了部分处理能力，但通过简化算法或减少 DSP 数量，仍能实现显著的时延降低。

2. **适用场景**：100m–2km 的中短距互联，适用于需要一定传输距离但仍对功耗敏感的场景。

### 3. 核心方案对比：

| 方案类型                                 | 技术原理                          | 时延表现                | 功耗优势             |
|--------------------------------------|-------------------------------|---------------------|------------------|
| <b>LPO (Linear Pluggable Optics)</b> | 完全移除 DSP 芯片，使用线性驱动器直接驱动光器件    | 极致低时延（无 DSP 处理延迟）   | 最高（功耗降低）         |
| <b>Hybrid / LRO</b>                  | 保留部分 DSP 功能 或简化 DSP 算法，结合线性驱动 | 超低时延（DSP 数量减半，时延减半） | 显著降低（相比全 DSP 0%） |
| <b>Full DSP</b>                      | 完整 DSP 芯片进行复杂均衡与纠错            | 较高（存在处理延迟）          | 较高（占模块 ~50%）     |

4. 易飞扬的半 DSP /Hybrid 光模块和有源光缆产品线是这个领域的工程创新型产品。比 LPO/LRO 具有超大型组网的现实可行性。

此类产品列表如下：



### Hybrid 单模光模块

| 产品类型                       | 波长     | 激光发射器 (LD) | 光电探测器 (PD) | 最大传输距离/长度 | 光接口形式  | 最大功耗 |
|----------------------------|--------|------------|------------|-----------|--------|------|
| 800G OSFP-HRO 2xDR4 光模块    | 1310nm | SiPh       | PIN        | 500m      | 双路 MPO | 13W  |
| 1.6T OSFP224-HRO 2xDR4 光模块 | 1310nm | SiPh       | PIN        | 500m      | 双路 MPO | 21W  |

## Hybrid 多模和单模 AOC

| 产品类型                          | 波长     | 激光发射器 (LD) | 光电探测器 (PD) | 最大传输距离/长度 | 光接口形式 | 最大功耗  |
|-------------------------------|--------|------------|------------|-----------|-------|-------|
| 400G QSFP-DD Hybrid VR8 有源光缆  | 850nm  | VCSEL      | PIN        | 50m       | AOC   | 6.6W  |
| 800G OSFP Hybrid VR8 有源光缆     | 850nm  | VCSEL      | PIN        | 50m       | AOC   | 9W    |
| 800G OSFP Hybrid PSM8 有源光缆    | 1310nm | SiPh       | PIN        | 500m      | AOC   | 12.5W |
| 1.6T OSFP224 Hybrid PSM8 有源光缆 | 1310nm | SiPh       | PIN        | 500m      | AOC   | 21W   |

## Hybrid 铜缆

| 产品类型                               | 波长  | 激光发射器 (LD) | 光电探测器 (PD) | 最大传输距离/长度 | 光接口形式 | 最大功耗 |
|------------------------------------|-----|------------|------------|-----------|-------|------|
| 800G OSFP Hybrid 增强型有源铜缆 (ACC+)    | 不适用 | 不适用        | 不适用        | 6m+       | 不适用   | 7W   |
| 1.6T OSFP224 Hybrid 增强型有源铜缆 (ACC+) | 不适用 | 不适用        | 不适用        | 4m        | 不适用   | 13W  |

## 二、采用 O-BAND 非相干 DWDM 作为 DCI 相干通信的有力补充，实现低时延低成本中短距 DCI

随着 AI 算力集群和分布式数据中心快速发展，数据中心互联 (DCI) 对 400G/800G/1.6T 高速光互联提出了更高要求。传统相干通信方案虽然具备超长距传输能力，但其复杂 DSP 架构、高功耗、高成本以及较高传输时延，在 30km 以内的中短距 DCI 场景中逐渐显现出成本与效率的不匹配。



基于此，易飞扬（GIGALIGHT）推出基于 O-BAND（1310nm 窗口）的非相干 DWDM 系列产品，通过利用 O-BAND 天然零色散特性，实现无需 DCM 色散补偿的高效传输架构，为数据中心中短距互联提供了可替代相干通信的高性价比解决方案。

相较于传统 C-BAND 相干方案，O-BAND 非相干 DWDM 方案具备以下核心优势：

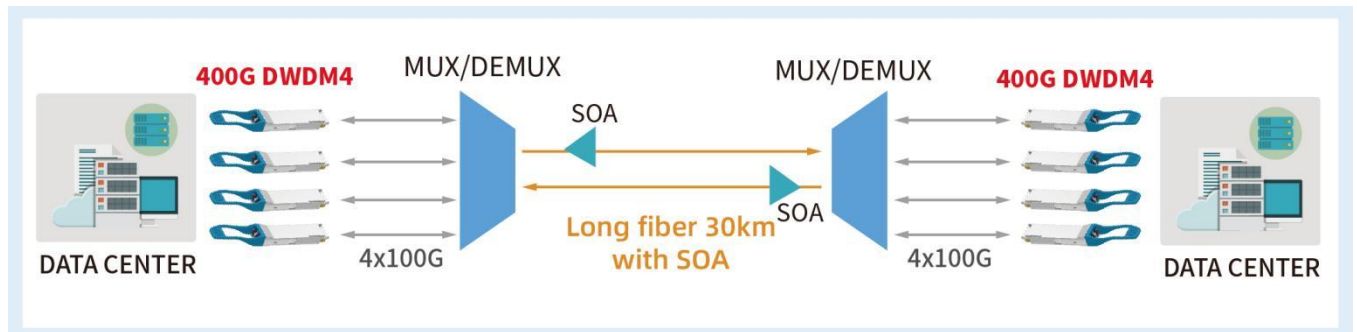
- **低时延**：省去复杂 DSP 色散补偿处理，显著降低链路传输延时，更适合 AI 训练、实时算力调度等低时延场景；
- **低功耗**：采用硅光集成技术与 PAM4 调制架构，整体功耗远低于传统相干模块；
- **低成本**：无需昂贵相干器件、DCM 模块及复杂线路设计，大幅降低 CAPEX 与运维成本；
- **高密度部署**：支持 QSFP-DD、OSFP 等高密度封装，满足下一代数据中心规模化部署需求；
- **灵活开放**：支持开放光网络架构，适配白盒交换机与开放线路系统，提升网络灵活性。

从 100G 到 800G，再到 1600G，易飞扬已经构建起业内最完整的 O-Band DWDM 可插拔产品矩阵，全面覆盖不同代际、不同容量的数据中心互联需求。

1. **100G QSFP28 DWDM1**：单波 100G，极简运维，系统支持 16 波传输；
2. **200G/400G QSFP-DD DWDM4**：4 路波长聚合，单波 50G/100G，系统支持 16 波传输；
3. **800G OSFP/QSFP-DD 2XPSM DWDM4 (O-BAND，单波, 100G PAM4)**系统支持 16 波传输；
4. **1.6T OSFP224-EL 2XPSM DWDM4 (O-BAND ,单波 200G PAM4,外置光源)**，系统支持 16 波传输。



## O-band DCI 互联应用场景 ( 2km-30km ) :



## 三、1.6T AI&DC 互连的三种灵活解决方案

随着 AI 算力需求爆发式增长，数据中心互联带宽正从 400G/800G 快速向 1.6T 演进。为满足不同场景下的部署需求，业界围绕 1.6T 形成了三种主流解决方案路径，各具特色，适配不同的成本结构、功耗预算与规模化诉求。

### 3.1 标准 1.6T 可插拔光模块 ( Pluggable Transceiver )

产品形态：标准 1.6T OSFP224 光模块

1.6T OSFP224 光模块是总带宽 1.6Tbps、采用 OSFP224 封装的新一代高速光模块，是当前满足 AI 集群和超算带宽需求的主流产品。

#### 核心参数与定义:

**总速率**：1.6 Tbps，通过 8 个 200G PAM4 通道实现总带宽，净载荷扣除开销后刚好为 1.6Tbps

**SFP224 封装**：OSFP 封装的演进版本，具备 224 个引脚，支持 16 路电气通道，相比传统封装密度更高、散热能力更强，适配高密度 AI 交换机部署。



| 产品类型             | 传输方案                      | 传输距离  | 核心优势                | 适用场景            |
|------------------|---------------------------|-------|---------------------|-----------------|
| 1600G OSFP DR8   | 8 路并行单模光纤，每路 200G         | 500m  | 方案成熟、适配短距互联         | AI 集群机架间/行内短距互联 |
| 1600G OSFP 2×FR4 | 两个 FR4 通道，每个通道采用 4 波长波分复用 | 约 2km | 节省约 75% 光纤资源，布线更加灵活 | 稍长距离的数据中心互联     |

### 3.2 外置光源的 1.6T 可插拔光模块 ( External Light Source Pluggable , ELS-Pluggable )

该方案将激光光源从 1.6T OSFP224 光模块中剥离，由机架级或板卡级集中式外置光源（ELS / Co-packaged Light Source）统一供光，通过光纤耦合至各插拔模块。模块本身只保留调制器、探测器及必要的电处理单元，大幅降低单模块热耗散。

作为外置光源 1.6T 可插拔光模块概念提出者，易飞扬可以提供如下外置光源可插拔解决

| 产品类型                       | 传输方案                      | 传输距离  | 核心优势           | 适用场景            |
|----------------------------|---------------------------|-------|----------------|-----------------|
| <b>1600G OSFP DR8 EL</b>   | 8 路并行单模光纤，每路 200G         | 500m  | 光源外置，大幅降低光模块功耗 | AI 集群机架间/行内短距互联 |
| <b>1600G OSFP 2×FR4 EL</b> | 两个 FR4 通道，每个通道采用 4 波长波分复用 | 约 2km | 光源外置，大幅降低光模块功耗 | 稍长距离的数据中心互联     |

### 3.3 外置光源的 1.6T ( 16×100G ) NPO 解决方案 ( Near-Package Optics , NPO )

NPO 方案是可插拔与共封装光学 ( CPO ) 之间的"过渡态"架构。光引擎 ( Optical Engine ) 不再封装于交换芯片内部，而是紧贴封装体外侧安装，配合外置光源通过硅光子或聚合物波导实现超短距高速互联。1.6T NPO 方案通常与硅光子 ( Silicon Photonics ) 平台深度结合，实现极低电-光传输损耗。必须注意，可实用的 NPO 光引擎基本必须基于 16×100G 或 32×100G 的物理架构。这是和 1.6T 可插拔光模块主要的界分点。虽然业界正在投票基于 224G SERDES 的 3.2T/6.4T NPO 架构，但这一标准尚未出台。

### 3.4 三种方案对照比较

| 对比维度            | 标准 1.6T 可插拔模块         | 外置光源 1.6T 可插拔模块 | 外置光源 1.6T NPO 方案       |
|-----------------|-----------------------|-----------------|------------------------|
| 光源位置            | 模块内部集成                | 机架/板卡级外置        | 机架/板卡级外置               |
| 光引擎位置           | 模块内                   | 模块内             | 紧邻 ASIC 封装外侧           |
| 功耗效率 ( pJ/bit ) | 较高                    | 中等              | 低 ( 最优 )               |
| 端口密度            | 中                     | 中               | 高                      |
| 与现有基础设施兼容性      | ✓ 高 ( 即插即用 )          | ✓ 较高 ( 端口兼容 )   | ✗ 低 ( 需重新设计主板/机架 )     |
| 运维与更换便利性        | ✓ 最佳                  | ✓ 较好            | △ 复杂 ( 需专业操作 )         |
| 热管理难度           | 高 ( 模块自发热 )           | 中 ( 光源集中管理 )    | 低 ( 热源分离且效率高 )         |
| 传输距离适用性         | 中短距 ( ≤2km )          | 中短距 ( ≤2km )    | 500m to 2km            |
| 系统架构复杂度         | 低                     | 中               | 高                      |
| 标准化成熟度          | ✓ 成熟 ( OIF/IEEE 已定义 ) | △ 推进中           | ✓ 成熟 ( 行业 OIF 协议可以支持 ) |

三种方案并非相互替代，而是针对不同阶段和场景的**分层部署策略**：

- 近期 ( 2024 ~ 2026 年 ) ：标准可插拔模块仍是主流，满足绝大多数数据中心的平滑升级需求；
- 中期 ( 2027 年+ ) ：外置光源可插拔方案很有可能将在超大规模 AI 集群中快速渗透，兼顾部

署便利与功耗优化；

- 远期（2027 年+）：NPO 方案随着硅光子与标准化成熟，有望成为 AI 超算互联的主导架构，为迈向 3.2T/6.4T 奠定基础。

## 四、采用 AWGR+EDFA/AWGR+SOA 实现关键节点枢纽的全光低时延交换

### 4.1 AWGR（阵列波导光栅路由器）

AWGR 作为无源波长路由引擎，利用其循环路由特性，将 WDM 信号中不同波长的光确定性地映射至对应输出端口，无需查表、调度或缓存，彻底消除了传统电交换的“光—电—光”转换环节。信号全程保持光域形态，单级交换时延仅为纳秒量级，较传统 OEO 架构降低 2~3 个数量级，且时延抖动极小，具备高度确定性，可满足算力低时延切换路由的业务需求。

### 4.2 AI 算力集群 AWGR+EDFA 拓扑结构

#### 1. 层级架构说明

该架构多用于数据中心跨机柜/跨园区长距离全光交换枢纽，拓扑采用“计算接入-核心波长交换-光放大输出”三级扁平结构：

- 接入层：每个 AI 计算节点/GPU 服务器通过 100G PSM DWDM4（C-BAND）光模块输出信号，每个端口直接携带 4 个 C 波段固定波长，信号通过光纤汇集到 DWDM 合波器（MUX）。
- 核心交换层：合波后信号送入  $N \times N$  AWGR 波长路由器，基于波长完成全光路由交换，输出侧信号完成光路切换。

- 补偿输出层：AWGR 交换后存在约 8dB 固有插损，经过 EDFA 光放大器放大补偿损耗后，输出到下一级交换节点或目的计算集群，满足 10km 级传输要求。

## 2. 拓扑特性

无多层 O-E-O 转换，核心交换时延 $<1\mu\text{s}$ ，整体架构简洁，适合拓扑相对稳定的大型算力枢纽部署。

技术参数对比：

| 参数        | 传统电交换 (OEO)                 | AWGR + EDFA 全光交换      |
|-----------|-----------------------------|-----------------------|
| 交换时延      | 500 ns ~ 数 $\mu\text{s}$    | $< 10\text{ ns}$ (单级) |
| 时延抖动      | 数 $\mu\text{s}$ ~ ms (队列波动) | $< 1\text{ ps}$ (热控后) |
| OEO 转换    | 每端口 1~2 次                   | 0 次 (全光域)             |
| 单端口功耗     | 15~25W / 100G               | 0.1~0.5W (EDFA 分摊)    |
| 1000 端口功耗 | 15~25 kW                    | ~455 W (80 信道示例)      |
| 协议依赖      | 需解析 IP/MPLS 头               | 完全透明，无需感知             |
| 缓存排队      | 存在，引入不确定时延                  | 无，波长路由器               |
| 信号格式透明    | 不透明，格式相关                    | 透明 (OOK/QPSK/QAM)     |

## 4.3 AI 算力集群 AWGR+SOA 拓扑结构

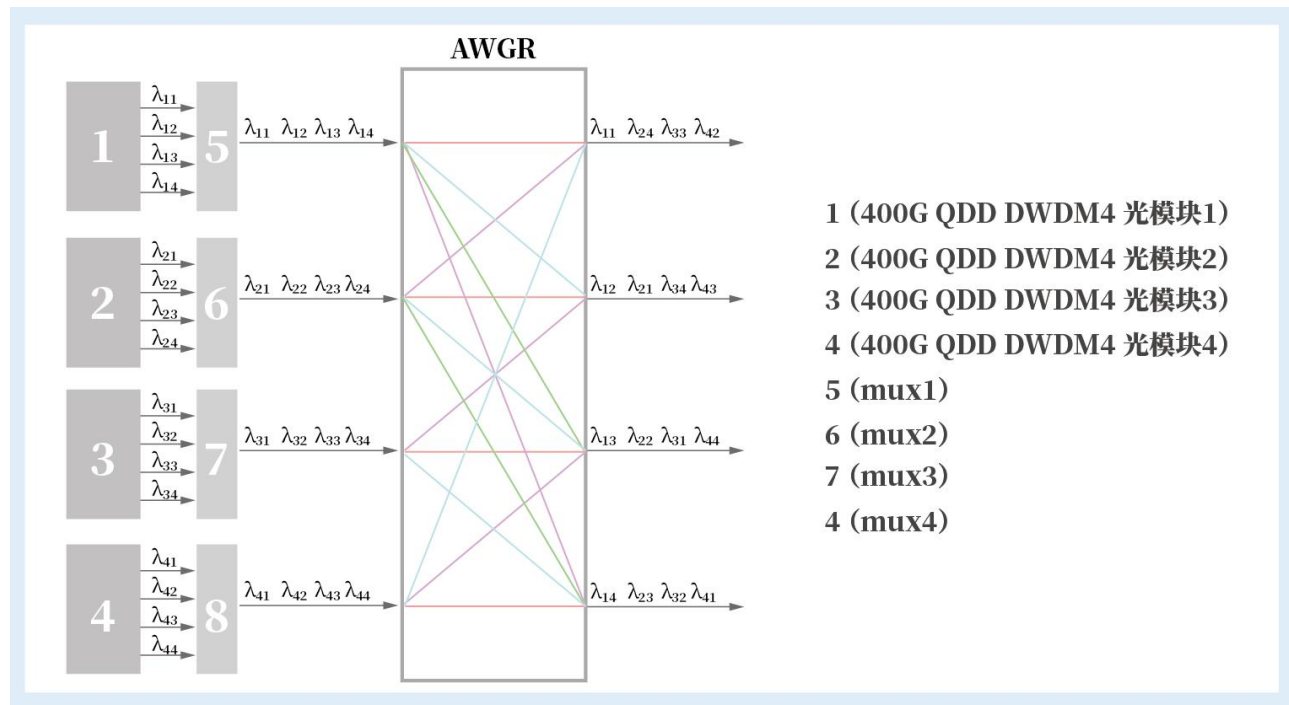
该架构多用于机柜内部/同机房短距离高密度全光交换枢纽，拓扑采用“有源调度+全光交换+集成放



大"紧凑结构：

- 1. 波长调度层：**每个计算节点连接 O-BAND DWDM 光器件或 O-BAND 多波长可调谐光引擎，调度控制器根据 GPU 通信需求（梯度同步/参数广播）动态调整输出波长。易飞扬提供不同波特率的 O-BAND DWDM 平行光模块适配中短距 AI 算力互连场景。
- 2. 核心交换层：**动态波长信号送入 AWGR 完成基于波长的全光路由，实现纳秒级光路切换。
- 3. 集成放大层：**采用集成化 SOA 半导体光放大器对交换后弱信号进行放大，相比 EDFA 体积更小，适合高密度机柜内部署，同时可辅助完成波长变换，提升调度灵活性。
- 4. 拓扑特性：**整个交换过程时延可压缩至 100 纳秒以内，单芯片功耗仅为同类电交换 ASIC 的 1/10，适合万卡级 AI 训练集群的 GPU 横向低时延通信。

组网示意框图如下：



#### 4.4 关键节点选型建议：



| 维度   | AWGR + EDFA            | AWGR + SOA               |
|------|------------------------|--------------------------|
| 适用场景 | 骨干网核心枢纽、大规模智算中心机架间长距互联 | 数据中心内部、高频交易、AI 集群短距低时延互连 |
| 交换粒度 | 波长级（粗粒度），微秒级切换         | 波长/时隙级（细粒度），亚纳秒级切换       |
| 时延表现 | 固定物理时延 + 微秒级重配置开销      | 固定物理时延 + 亚纳秒级重配置开销       |
| 信号质量 | 高（低噪声，适合长距）            | 中（需考虑非线性与噪声）             |

## 结束语

由于算力网的本质是工具调用的能力。单纯的算力网——这一昂贵的基础设施，如果没有全民好学的基础，算力网将是没有意义的。从任何商业的发展角度攀附算力大模型，这将和语言为建构的知识事业发生冲突。商人组织只能在算力网上思考创新，他们不能希望算力本身是竞争能力。算力网必须是惠民智力的网络。否则花费万亿巨资产生不了任何正向的效果。正如以往的电信管道是连接信息的目的。算力网必须是连接人的工具智能的目的。只有对算力网的目的做好精确的竞争力的定义，算力网将成为我们国家愿意埋单的一张网络。