

# Consideration on NPO and CPO at 400G/Lane

Runlong Hu (China Mobile)

Haojie Wang (China Mobile)

Weiqiang Cheng (China Mobile)

# Overview

- Requirement
- Technical Benefit
- Optics and FEC consideration
- Summary

# Requirements for NPO/CPO in AI Clusters

- Scaling AI clusters encounter power crisis
  - A 200,000-GPU cluster consumes 435 MW of IT<sup>[1]</sup> power; optical transceivers alone account for 17 MW <sup>[2]</sup>.
- CPO (Co-Packaged Optics) and NPO (Near Package Optics) are among the technologies to alleviate the crisis
  - CPO integrates optics directly into the same substrate as the ASICs<sup>[6]</sup>; NPO positions optical components close to the ASIC packages, but not on the same substrate<sup>[6]</sup>.
  - Compared to traditional DSP-based 800G pluggables with ~15W, CPO cuts this to ~5W (Meta @ ECOC 2025<sup>[2]</sup>, ASE 2025<sup>[3]</sup>) by removing the power-hungry DSPs.
  - Compared to liner pluggable optics (LPO, removes the DSP/retimer), CPO/NPO shorten the electrical path from tens of centimeters to millimeters, which reduces the difficulty of high-speed signaling integrity, especially at 400G/lane.
- The industry has already launched solutions, and application deployment has also begun.
  - NVIDIA CPO switches (Quantum-X / Spectrum-X)<sup>[4]</sup>, Broadcom Tomahawk / Jericho families 200G/lane CPO <sup>[5]</sup>.
  - In China, multiple switch manufacturers have also successively released related products.
  - Some hyperscalers have committed to deployment roadmaps:
    - Meta is driving 800G CPO pilot programs to replace pluggables in spine-leaf AI fabrics<sup>[2]</sup>
    - China's industry is expected to begin gradual deployment from 2027 to 2029.

[1] IT power (or IT load): the power consumed exclusively by the actual computing equipment

[2] <https://newsletter.semianalysis.com/p/co-packaged-optics-cpo-book-scaling>

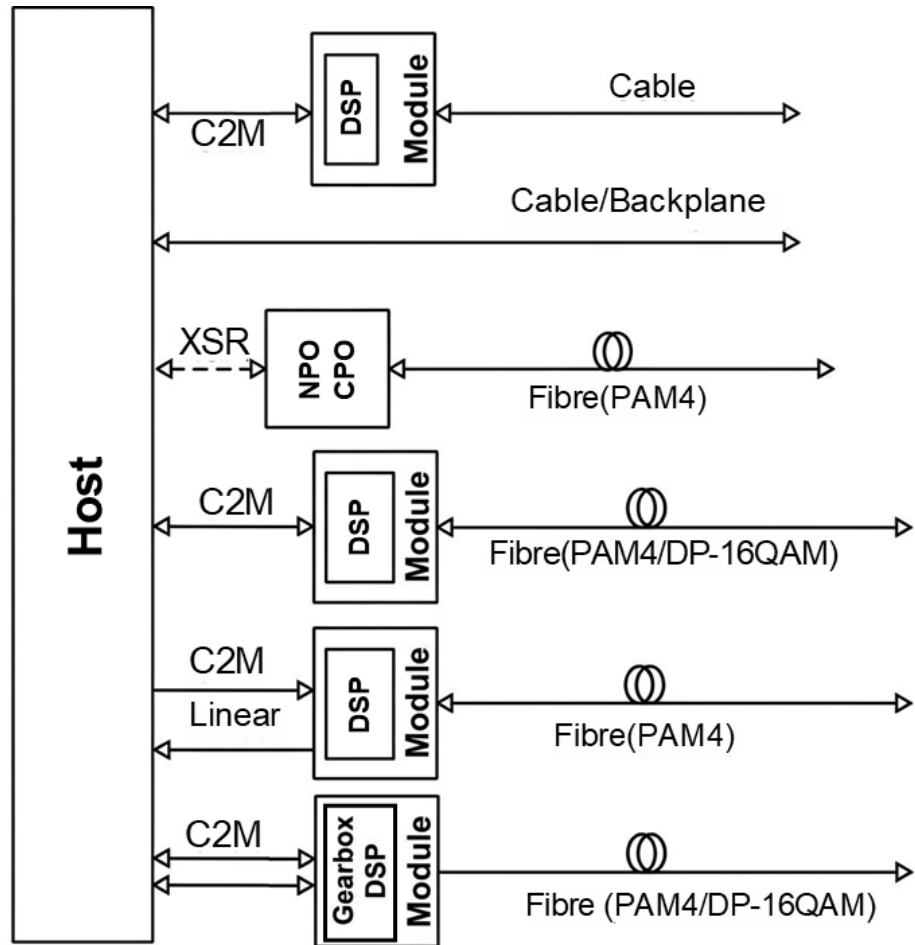
[3] <https://ase.aseglobal.com/press-room/ase-demonstrates-cpo-for-ai-applications>

[4] <https://www.nvidia.com/en-us/networking/products/silicon-photonics/>

[5] <https://investors.broadcom.com/news-releases/news-release-details/broadcom-announces-third-generation-co-packaged-optics-cpo>

[6] S. Priyadarshi, "Unlocking the Potential of Co-Packaged Optics in AI and HPC: Opportunities and Challenges," in IEEE Communications Magazine, vol. 64, no. 2, pp. 24-29, February 2026

# Technical Benefit — From an architectural perspective



- **Benefits of NPO/CPO**

- Shorter electrical path → lower insertion loss → better signal integrity
- Power savings from reduced equalization and driver strength
- Enables higher port density for AI infrastructure

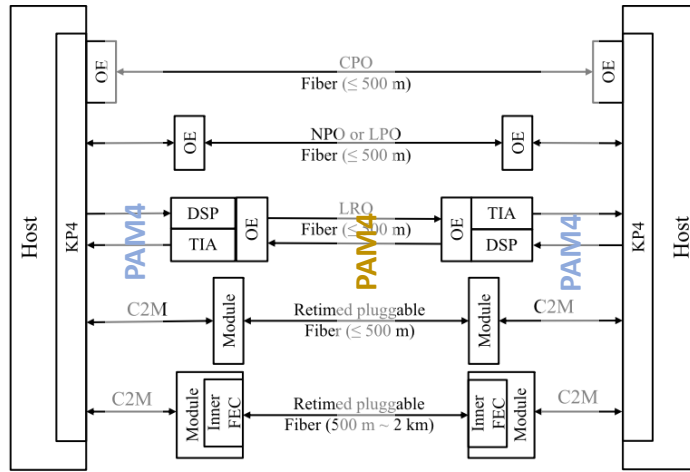
- **The trade-off**

- NPO/CPO achieves this by eliminating the retimer/DSP entirely
- The retimer is not merely an amplifier — it is the **translation layer** between electrical and optical domains
- It handles: rate adaptation, gearboxing, FEC alignment, and modulation mapping
- Without it, electrical and optical domains must align at the same rate and modulation format

# Optical Constraint

- The optical industry has converged on PAM4 for 400G/lane, optical modulators (InP, SiPh, TFLN) are designed and characterized for PAM4 with above 100 GHz bandwidth
  - InP modulators: demonstrations beyond 160 GBd PAM4
  - Silicon Photonics: 90+ GHz bandwidth MZM demonstrated; IMEC reported 110 GHz GeSi EAM for 400G PAM4
  - TFLN: demonstrations beyond 110 GHz, viewed as a major enabler for high-bandwidth modulation
- The optical modulation forces a firm constraint that the electrical side needs to accommodate
  - Electrical PAM6 baud rate does not equal optical PAM4 baud rate:
    - PAM4: 2 bits/symbol. At 400G/lane: ~212.5 GBaud, Nyquist ~106.25 GHz;
    - PAM6: ~2.585 bits/symbol (5 bits → 2 symbols via QAM-32 mapping). At 400G/lane: ~170 GBaud with KP4, Nyquist ~85 GHz
  - In pluggable modules, the retimer/gearbox handles this translation; In NPO/CPO, however, there is no retimer — electrical and optical must run at the same rate

# FEC Dimension



potential FEC architecture for PAM4

- **SNR degradation**

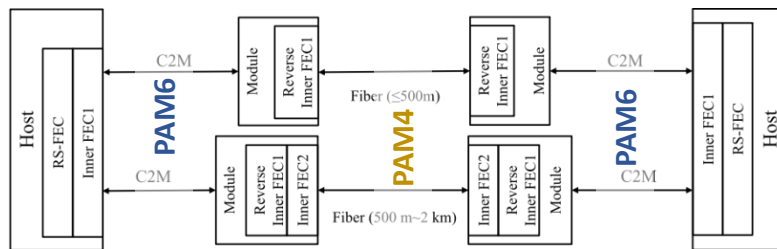
- PAM6 has degraded SNR compared to PAM4 for the same peak

- **FEC implications**

- Degraded SNR requires stronger FEC to maintain target BER
  - Stronger FEC typically means concatenated codes: outer code (KP4) + **inner code**
  - The inner code is conventionally processed by the retimer / gearbox / DSP module
  - PAM4 can use a lightweight inner FEC or even KP4 alone, while PAM6 requires stronger inner FEC

- **The NPO/CPO barrier:**

- In NPO/CPO, there is no retimer to host the inner code
  - Therefore, accommodating PAM6's FEC requirements in NPO/CPO becomes significantly more challenging



potential FEC architecture for PAM6

# Summary and Recommendations

- This contribution provides some thoughts on considering CPO and NPO for 400G/lane from the perspectives of architecture, optical technology, and FEC.
- Given the significant power reduction advantages brought by CPO and NPO, the industry will certainly spare no effort to promote their application, especially at 400G/lane.
- Incorporating CPO/NPO into the standardization work for 400G/lane would be encouraging to the industry and help reduce the difficulty of subsequent application.

Thanks