



## 주요 동향(2) : ICT

### 1 AI 데이터센터의 데이터 전송 병목 해소를 위한 광통신 전환

⇒ AI 데이터센터 병목, 연산에서 네트워크로 이동

- GPU 성능 경쟁에서 클러스터 연결 효율 경쟁으로의 전환
  - AI 모델의 규모가 빠르게 커지는 동시에, 이를 실시간으로 구동하는 추론 수요까지 폭증하면서 데이터센터에 요구되는 연산 처리량이 급격히 확대
  - 이에 따라 수천~수만 개 GPU를 고속 네트워크로 묶어 하나의 거대한 컴퓨터처럼 운영하는 클러스터 방식이 AI 인프라의 기본 단위로 정착
  - 클러스터 환경에서는 개별 GPU의 연산 성능보다 GPU 간 데이터 교환 속도가 전체 성능을 좌우하게 되면서, 인프라 경쟁의 축도 칩 성능에서 칩 간 연결 효율로 이동
- 데이터 이동 부담 확대와 기존 연결 구조의 한계
  - 클러스터 환경에서는 스위치를 통과하는 데이터 총량 자체가 급증하면서, 데이터를 더 적은 전력으로, 안정적으로 이동시키는 것이 AI 인프라의 핵심 과제로 부상
  - 현재 구조에서는 칩에서 생성된 전기 신호가 회로기판 위를 15~30cm 이동한 뒤 전면 패널의 플러거블 광모듈에서 빛으로 변환되는데, 이처럼 긴 전기 구간에서 신호 손실과 전력 소모가 발생
  - 현재 사용되는 800G급 플러거블 광모듈은 개당 약 14~18W의 전력을 소비하며, 스위치 1대에 64개 포트가 장착될 경우 광모듈만으로도 약 900W 이상의 전력이 소모
  - 이 같은 부품 단위 전력 부담이 누적되면서 네트워크 전력은 데이터센터 전체 소비 전력의 10%를 상회하고 있으며, 클러스터 규모 확대에 따라 그 비중도 지속적으로 상승
- 전기에서 광 중심으로, 데이터 이동 방식의 광전환과 CPO의 부상
  - 문제는 칩에서 생성된 전기 신호를 구리 배선 위로 전달하는 기존 전기 기반 연결 구조만으로는, 확대되는 데이터 이동 부담과 차세대 대역폭 요구를 더 이상 감당하는 것이 어려워지고 있다는 점

- 특히 전송 속도가 빨라질수록 구리 구간의 신호 손실이 커지고 이를 보정하는 칩(DSP)도 더 많은 전력을 쓰게 되면서 속도와 전력 효율을 함께 높이기 어려운 구조가 심화
- 이에 따라 데이터센터 인터커넥트는 단순한 부품 교체를 넘어, 데이터를 이동시키는 방식을 전기 중심 장거리 전달 구조에서 광 중심 근접 변환 구조로 전환하는 방향으로 재편
- 구체적으로는, 기존에 스위치 전면 패널에 분리 장착되어 있던 광 변환 기능을 스위치 패키지 안으로 통합하여 전기 신호가 이동해야 하는 거리 자체를 수 mm 수준으로 줄이는 접근이 핵심
- 이 방식을 구현하는 기술이 CPO(Co-Packaged Optics)로, 광학 부품과 연산 칩을 하나의 패키지에 함께 설계하는 차세대 인터커넥트 기술
- 이러한 전환 속에서 CPO는 단순한 부품 배치 변경을 넘어, 데이터 이동의 병목을 완화하고 차세대 스위치 및 대규모 클러스터의 연결 효율을 뒷받침하는 핵심 대안으로 부상
- '25~'26년 들어 NVIDIA·Broadcom·TSMC가 각각 시스템·반도체·제조 영역에서 CPO 상용화를 본격 추진하면서, 광전환은 기술 검증 단계를 넘어 양산 준비 국면으로 진입

➔ AI 데이터센터의 광전환 경로와 적용 단계

● 광모듈 분리형 구조에서 패키지 통합형 설계 구조로의 전환

- 기존 플러거블 방식에서는 광모듈과 연산 칩이 물리적으로 분리돼 각각 독립적으로 설계·제조·교체할 수 있었으나, CPO에서는 광 엔진이 칩과 같은 패키지 수준에서 통합되며 하나의 구조로 결합
- 이 통합은 단순한 부품 위치 변경이 아니라, 칩 설계 단계부터 광 엔진의 배치와 열 분포, 신호 경로를 함께 반영해야 하는 변화로, 제조·패키징·테스트·유지보수 방식 전반의 재조정을 수반
- 이 때문에 CPO는 기존 광모듈을 더 빠른 제품으로 교체하는 수준의 개선이 아니라, AI 칩 패키지의 설계 원칙 자체를 광 중심으로 재편하는 아키텍처 전환으로 평가

● 전력·거리 부담이 큰 구간부터 시작되는 단계적 광전환

- 데이터센터의 모든 연결을 한꺼번에 광으로 전환하기보다, 연결 거리가 길고 전력 부담이 큰 구간부터 광 기반 기술을 단계적으로 도입하는 방식이 현실적인 경로로 자리매김



- 같은 랙 안에서 GPU끼리 연결하는 0.5~2m 내외의 짧은 구간에서는 당분간 구리 직결 방식이 비용 측면에서 유리하고 별도 광 변환이 필요 없어 링크 전력 부담도 상대적으로 작아서 여전히 유리
- 이에 따라 데이터센터 내부에서는 연결 길이가 짧은 구간은 구리, 랙 간·클러스터 간과 같은 장거리 구간은 광이 담당하는 병행 구조가 형성
- 이 중 스위치는 수십~수백 개의 고속 광 포트가 집중되는 네트워크의 교차 지점으로, 데이터센터 내 전력과 대역폭 부담이 가장 커 CPO를 가장 먼저 적용하기 유리한 장비
  - (전력 집중도) 포트 수가 많을수록 광모듈이 차지하는 전력 비중이 함께 커지기 때문에, CPO에 따른 절감 효과가 다른 장비보다 스위치에서 가장 크게 나타나는 구조
  - (열 환경) 스위치는 GPU보다 상대적으로 안정적인 열 조건을 갖추고 있어, 온도 변화에 민감한 광소자의 성능을 유지하는 데 더 유리한 환경을 제공
  - (효과성) 반면 GPU는 상대적으로 짧은 구리 기반 전기 연결 비중이 높고, 스위치처럼 많은 광 포트가 한 장비에 집중되는 구조가 아니라서 같은 방식의 효과가 상대적으로 제한적인 단계
- 스위치에 CPO를 적용하면 광모듈 전력 감소를 중심으로 스위치 장비의 전체 전력을 의미 있게 줄일 수 있고, 전력 감소는 장비 발열 저감으로 냉각 설비 부담과 관련 에너지 비용도 완화
- 이러한 흐름 속에서 광전환은 기존 플러거블 광모듈의 개선부터 광 엔진의 근접 배치, 연산 칩 직접 통합(Optical I/O)으로 이어지는 여러 기술 단계를 중심으로 전개
- 외부 레이저(ELS) 구조를 통한 유지보수 부담 완화
  - 스위치가 CPO의 첫 적용처가 되는 이유는 전력 효율뿐 아니라, CPO 도입 시 가장 큰 운영 우려인 유지보수 문제도 스위치 구조에서 가장 자연스럽게 해소할 수 있기 때문
  - CPO의 가장 큰 운영상 우려는 광 엔진이 칩 패키지에 통합되면서 고장 시 광모듈만 교체하던 방식이 어려워진다는 점이며, 특히 열에 민감하고 수명이 제한된 레이저가 핵심 부담 요인으로 지목
  - 이에 따라 업계는 레이저를 패키지 밖 별도 모듈에 두고 광섬유로 빛을 공급하는 ELS 방식을 채택하면서, 레이저 고장 시 스위치 전체가 아니라 레이저 모듈만 교체할 수 있는 구조를 마련

➔ 글로벌 상용화 경쟁과 광전환의 다음 단계

1. CPO 상용화를 이끄는 시스템·네트워크·제조 역할 분화

● 역할 분담 기반의 CPO 상용화 생태계

- CPO 상용화는 단일 기업이 독자적으로 완성할 수 있는 영역이 아니라, 수요 창출, 기술 실증, 양산 제조를 담당하는 서로 다른 기업들이 함께 움직여야 가능한 생태계
- 이 생태계는 AI 클러스터 수요를 이끄는 시스템 사업자, CPO 기술을 실증하는 네트워크 반도체 사업자, 이를 양산 가능한 형태로 구현하는 제조 플랫폼 사업자의 세 축으로 구성
- 현재 이 세 축을 대표하는 기업이 각각 NVIDIA, Broadcom, TSMC이며, 이들의 전략과 진척 속도가 CPO 상용화 흐름에 가장 큰 영향을 미치고 있는 구도

● (NVIDIA) AI 클러스터 수요를 견인하며 광 공급망까지 직접 확보

- NVIDIA는 GPU를 넘어 스위치, 네트워크, 소프트웨어까지 아우르며 AI 인프라 전체를 설계하는 시스템 사업자로, CPO를 자사 클러스터 아키텍처의 필수 구성 요소로 규정
- AI 데이터센터에 주로 사용하는 두 가지 네트워크 방식\* 모두에 맞는 CPO 스위치를 출시해, 네트워크 종류와 관계없이 광전환을 적용할 수 있는 제품 구성을 갖추
- NVIDIA CPO 스위치는 TSMC의 광반도체 기술(전자 칩과 광 칩을 함께 묶는 기술)을 적용해 기존 플러거블 방식보다 전력은 덜 쓰고 신호는 더 안정적으로 전달할 수 있도록 설계
- NVIDIA의 차별적 전략은 기술 확보에 머무르지 않고, 레이저·광 DSP·실리콘 포토닉스까지 포괄하는 광 공급망을 대규모 투자를 통해 직접 확보
- 이는 CPO 상용화 과정에서 핵심 부품의 조달 안정성과 기술 주도권을 동시에 확보하려는 시스템 사업자 특유의 수직적 접근

● (Broadcom) 스위치 ASIC-광 엔진 자체 설계 기반의 CPO 양산 실증

- Broadcom은 스위치 ASIC과 광 엔진을 모두 자체 설계하는 네트워크 반도체 기업으로, 업계에서 가장 이른 시점부터 CPO 개발에 착수해 기술 성숙도와 양산 경험을 축적
- 1세대(21년)에서는 광 엔진 통합 공정을 검증하고, 2세대(24년)에서는 양산형 CPO를 구현하며 기존 플러거블 대비 약 30% 이상의 시스템 전력 절감 효과를 실증



- 3세대('25년)에서는 한 채널이 처리하는 데이터 속도를 두 배로 높였고, 다음 단계에서는 이를 다시 두 배로 높이는 계획까지 제시하며 앞으로의 기술 발전 방향을 구체화
- 또한 광부품, 스위치, 연결 부품, 네트워크 장비를 맡는 주요 협력사들과 함께 양산 준비 현황을 공개하며, CPO 상용화에 필요한 생태계가 형성되고 있음을 보여줌
- (TSMC) COUPE 플랫폼으로 CPO 제조 기반 구축
  - TSMC는 CPO를 직접 설계하는 기업이 아니라, 이를 양산 가능한 형태로 구현하는 제조 플랫폼 사업자로서 NVIDIA와 Broadcom 양쪽에 공통 제조 기반을 제공
  - TSMC의 실리콘 포토닉스 패키징 플랫폼 COUPE는 전자 칩과 광 칩을 아주 가깝게 직접 붙여, 칩 사이를 오가는 전기 신호 거리를 줄이고 전력 손실도 낮추는 광반도체 패키징 기술
  - TSMC는 먼저 플래거블 광모듈에서 관련 생산기술을 시험한 뒤, 이를 CPO까지 확대 적용하는 생산 계획을 제시하며 단계적으로 제조 전환을 추진
  - COUPE는 TSMC의 파운드리 고객이 활용할 수 있는 개방형 제조 플랫폼으로, 대형 반도체 기업뿐 아니라 중소 칩 설계사까지 CPO 기술에 접근할 수 있는 경로를 제공
  - 실제로 반도체 설계 기업 Alchip과 광 통신 칩 기업 Ayar Labs가 COUPE 기반 광 I/O 서브 시스템을 시연해, 자체 광 기술이 없는 기업도 자사 가속기에 광 연결을 탑재할 수 있음을 입증

## 2. 스위치-CPO 이후, 연산 칩 패키지 내 광 통합(Optical I/O)으로의 확장

- GPU 측 연결 병목의 심화와 Optical I/O 도입 필요성
  - 스위치 CPO로 외부 연결 부담은 줄일 수 있지만, GPU끼리 직접 연결되는 내부 구간은 여전히 전기 방식이어서 클러스터가 커질수록 이 부분이 다음 병목으로 부상
  - 특히 클러스터가 수백~수천 GPU로 확대되면 개별 GPU가 교환해야 하는 데이터 총량이 급증하면서, 구리 기반 I/O 용량 자체가 한계에 근접하는 구간이 발생
  - 이 병목을 해소하려면 스위치뿐 아니라 GPU 패키지에도 광 인터페이스를 직접 통합하여, GPU 간 통신이 출발점부터 도착점까지 전 구간 광으로 이루어지는 구조로의 전환이 필요

- 이에 따라 주요 가속기 설계사들도 GPU 패키지 수준의 광 통합을 준비하기 시작하면서, 광전환의 범위가 스위치를 넘어 연산 칩 자체로 확대되는 흐름
  - GPU 패키지 광 통합의 난이도로 인한 적용 시차 확대
    - 스위치 CPO에서는 광 엔진과 스위치 ASIC을 하나의 패키지에 통합하는 것이 핵심 과제였다면, Optical I/O에서는 광 엔진이 GPU, HBM, SRAM과 하나의 패키지 안에 함께 집적되어야 하는 구조
    - 이는 칩 설계, 패키징, 메모리 계층 전반을 광 인터커넥트를 전제로 재설계해야 한다는 점에서, 스위치 CPO 대비 한 차원 높은 통합 복잡도를 수반
    - 이 복잡도 차이로 인해 스위치 CPO가 '26년 본격 상용화를 개시하는 반면, GPU 패키지 내 Optical I/O는 2~3년 후행해 '28~'29년 이후 실질적 적용이 시작될 것으로 전망
  - Optical I/O 실현에 따른 데이터센터 인터커넥트 구조 재편
    - Optical I/O가 GPU 패키지에 적용되면, GPU 내부 연결까지 광으로 바뀌면서 데이터센터 전체가 부분 광전환에서 전반적 광 연결 구조로 확장
    - 이에 따라 GPU와 스위치 사이에 남아있던 마지막 구리 구간이 제거되면서, 수천 개 가속기를 하나의 통합된 시스템처럼 운용할 수 있는 기반이 형성
    - 이는 기존에 클러스터 규모의 상한선을 결정하던 구리 연결 거리와 대역폭 제약이 해소된다는 의미로, 더 큰 규모의 AI 모델을 별도 클러스터로 분할하지 않고 단일 시스템에서 운용 가능
    - 또한 전 구간 광전환은 구리 연결에서 발생하던 전력 낭비까지 줄여, 데이터 이동의 에너지 효율을 한 단계 더 높이는 효과도 수반
- ### 3. CPO 상용화를 좌우하는 핵심 과제와 확산 조건
- (열·신뢰성) 고열 환경에서의 광소자 안정성과 레이저 신뢰성 확보
    - 연산 칩에서 발생하는 고열이 바로 옆의 광소자와 레이저에 직접 영향을 미치는 환경에서, 광 부품의 성능과 수명을 안정적으로 유지하는 것이 CPO의 핵심 기술 과제
    - 광소자는 1°C의 온도 변화만으로도 약 0.1nm의 파장 이동이 발생, 신호 품질이 저하될 수 있으며 특히 레이저는 열에 가장 민감한 부품으로 CPO의 대표적인 신뢰성 병목으로 지목
    - 한편, Meta가 OFC 2026에서 공개한 자료에서는 CPO가 플러거블 대비 더 높은 신뢰성을 보인다는 결과가 제시되어, 기계적 접촉 부품이 적은 CPO 구조의 이점이 데이터로 확인되기 시작



- 다만 이러한 검증 결과가 다양한 운영 환경과 장기간에 걸쳐 축적되어야 업계 전반의 채택과 확산으로 이어질 수 있어, 추가적인 현장 검증이 지속적으로 필요한 단계
- (제조·검증) 광·전자 이중 집적의 양산 수율과 테스트 체계 구축
  - CPO 대량 생산을 위해서는 광섬유 정렬, 통합 패키징, 수율 확보, 테스트 체계를 모두 갖춰야 하는데, 이 공정들이 개별적이 아닌 동시에 안정화되어야 한다는 점이 양산 난이도를 높이는 요인
  - (광섬유 정렬) 광섬유와 광 칩의 연결에는  $0.1\mu\text{m}$  수준의 정밀도가 요구되며, 이를 대량 양산 환경에서 자동으로 수행하는 기술이 아직 충분히 성숙하지 않은 상황
  - (광 테스트) 양산 수율을 확보하려면 광 엔진을 패키지에 조립하기 전에 개별 동작 여부를 미리 검증해야 하나, 이를 위한 전용 검사 장비와 측정 방법이 아직 개발 초기 단계에 위치
- (운영·표준화) 패키지 통합 이후 유지보수 방식과 표준 생태계 정립
  - 광 엔진이 패키지에 통합되면서 기존 플러거블처럼 모듈 단위로 교체하는 방식이 어려워지기 때문에, 외부 레이저(ELS) 관리 구조를 포함한 새로운 유지보수 체계의 확립이 필요
  - 현재 CPO 관련 표준은 아직 초기 단계여서, 광 엔진과 관리 시스템이 회사가 달라도 서로 호환될 수 있어야 여러 기업이 함께 참여하는 생태계로 확산 가능
  - 표준화가 충분히 진전되고 양산 인프라가 성숙해 비용이 낮아져야, CPO가 하이퍼스케일러 전용 기술을 넘어 데이터센터 전반의 범용 인터넥트로 확산될 수 있는 조건이 형성

출처 : Semiconductor Engineering 외(2026.3.)

<https://semiengineering.com/cpo-is-extending-the-limits-of-whats-possible-in-ai-data-centers/>  
<https://semiengineering.com/verifying-scale-up-and-scale-out-in-data-centers/>  
<https://stock.mk.co.kr/invest/view/3579>  
<https://semiengineering.com/co-packaged-optics-reaches-power-efficiency-tipping-point/>  
<https://www.taiwannews.com.tw/news/6291090>  
<https://investors.broadcom.com/news-releases/news-release-details/broadcom-announces-third-generation-co-packaged-optics-cpo>  
<https://www.taiwannews.com.tw/news/6291090>  
<https://pr.tsmc.com/english/news/3136>  
<https://www.tomshardware.com/tech-industry/semiconductors/industrys-first-tsmc-coupe-based-optical-connectivity-solution-for-next-gen-ai-chips-displayed-alchip-and-ayar-labs-show-future-silicon-photonics-device>