Optical Networking & AI Infrastructure

A sector primer: how light gets from a wafer of indium phosphide to a hyperscale AI cluster, and where the physical supply chain actually strains.

Executive Summary

Every dollar of AI infrastructure spending eventually runs through a piece of glass smaller than a grain of rice: an optical chip that turns electricity into pulses of light, and light back into electricity, fast enough and far enough to hold a cluster of GPUs together. This brief is not a call on any stock or any name. It is a plain explanation of how the optical networking sector physically works, from raw indium and phosphorus through to a transceiver plugged into a hyperscaler's switch, so that any view on a specific company can be built on the underlying mechanism rather than borrowed from someone else's price target.

Three things matter more than any single quarter's earnings print. First, optical connectivity stops being optional past a certain distance and speed: copper simply runs out of usable reach once GPU clusters spread across racks, rooms, and buildings. Second, the industry sits on top of a supply chain that narrows sharply the further upstream you look, from dozens of companies capable of assembling a transceiver down to two or three companies capable of making a laser chip that actually clears a major customer's qualification bar. Third, every generational speed jump, from 400G to 800G to 1.6T and now 3.2T, follows the same script: fast unit growth, and even faster price decay, which is why the dollar size of the market does not simply grow in a straight line even as demand for bandwidth clearly does.

THIS REPORT COVERS

Section 1	**Why Optics, Why Now**	The bandwidth wall that makes copper unworkable at AI cluster scale
Section 2	**The Value Chain**	From raw indium and phosphorus to a rack in a data center
Section 3	**Anatomy of a Transceiver**	What is actually inside the module, and what each part costs
Section 4	**The Laser Is the Hard Part**	EML versus CW-plus-silicon-photonics, and why the laser is the real chokepoint
Section 5	**Speed & Form Factor**	400G to 3.2T, and the slow march from pluggables toward co-packaged optics
Section 6	**The Upstream Squeeze**	Why indium phosphide, specifically, is the sector's binding constraint
Section 7	**Sizing the Market**	What the numbers say about growth, and about price decay
Section 8	**Where Informed Opinion Disagrees**	Three genuine open debates, and what each side is actually claiming

SECTION 1

Why Optics, Why Now

Every wire has a speed limit that gets worse with distance. Push enough electrical signal down a copper cable, fast enough, and past a certain length the signal degrades faster than it can be cleaned up. That limit used to be a non-issue for data centers, because the things that needed to talk to each other were close together. AI training clusters broke that assumption. A single training run now spans not just one rack of GPUs, but many racks, often many rooms, and increasingly many buildings.

Engineers in the field generally describe this in three distances, each with its own connectivity demands:

- **Scale-up:** the fastest, shortest connections, chip to chip and GPU to GPU inside a single server or rack.
- **Scale-out:** connecting racks of GPUs to each other inside the same data center, to form one large, coordinated compute cluster.
- **Scale-across:** the newest and fastest-growing category: connecting entire data centers to each other, so that a training workload can span more capacity than any single building can hold.

Copper still works well for scale-up, over very short distances, and it remains the cheapest option where it is physically usable. But as the distance grows, copper's signal loss becomes the limiting factor, while light traveling down a fiber loses very little signal strength over the same distance. That physical difference, not a preference or a trend, is the reason optical connectivity share inside data centers has been rising steadily. One senior industry veteran, formerly responsible for technology strategy at Lumentum, put a number on this shift directly: optical connectivity inside the data center was handled photonically at around 30% of total connections a few years ago, a figure he estimated has since risen toward roughly 60% for some architectures, before scale-across is even added on top.

This is the physical reason the optical sector exists as a distinct, fast-growing part of the AI buildout. Everything in the sections that follow, the chips, the lasers, the scarcity, the pricing, is downstream of this one constraint: past a certain point, only light does the job.

SECTION 2

The Value Chain: Raw Material to Rack

Trace an optical transceiver back to its origin and the chain is shorter, and more concentrated, than most people expect. Six stages sit between a shovel of raw material and a working AI cluster.

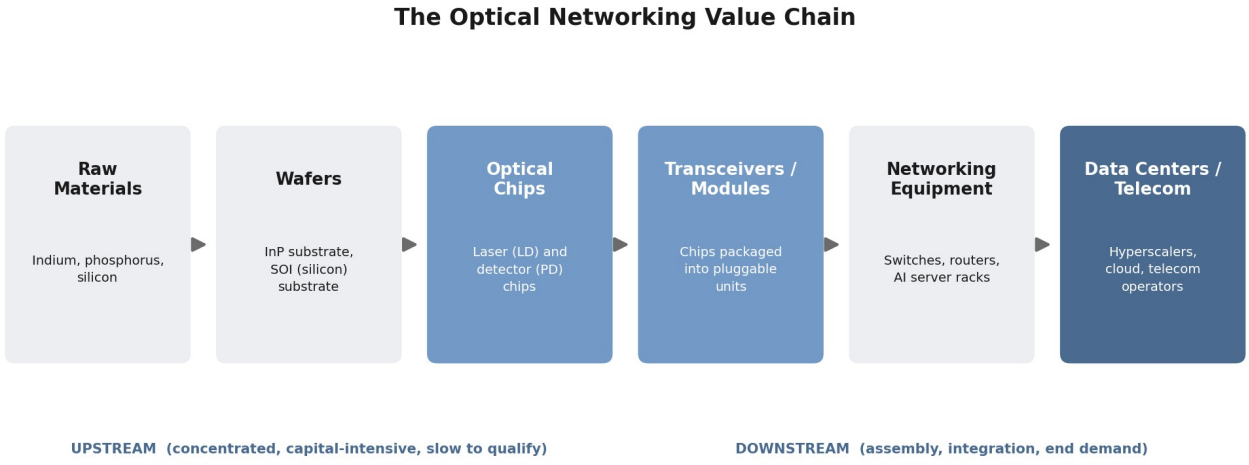

Figure 1. The optical networking value chain, from raw material to end demand.
Source: Nomura Research, company data, OsloMet Capital Group

The number of companies capable of operating at each stage shrinks sharply as you move upstream. Dozens of firms assemble transceivers and modules. A much smaller set can reliably manufacture the underlying optical chips. Fewer still can grow usable indium phosphide crystal at scale. This is not a coincidence: each step upstream requires more specialized, capital-intensive, slow-to-qualify manufacturing know-how than the step after it. It is also the single most important structural fact in the sector, because it means the companies furthest downstream, however strong their execution, are ultimately dependent on the handful of companies furthest upstream simply being able to keep them supplied.

It is worth being precise about where different types of companies actually sit in this chain, because sell-side coverage often blurs the distinction. Laser and chip makers (companies like

Lumentum, Coherent, and Applied Optoelectronics) sit at the chip and module layers. Contract manufacturers and ODMs (companies like Celestica) sit downstream of all of this, assembling the switches and server racks that the modules plug into; they are customers of the optical supply chain, not participants in it. Strong results from a systems assembler are a demand signal for the whole chain, not a direct read on chip-level scarcity.

SECTION 3

Anatomy of a Transceiver

A transceiver is the small pluggable unit, roughly the size of a cigarette lighter, that plugs into a network switch port. Its job is simple to state and hard to execute: convert an electrical signal into pulses of light for transmission down a fiber, and convert incoming light back into an electrical signal on the receive side. Every AI server rack needs many of these, and the faster the target speed, the more of them, and the more fibers and lasers, a single rack requires without the module itself getting any bigger.

Inside the housing, a transceiver splits into two mirrored halves. The TOSA, or Transmit Optical Sub-Assembly, contains the laser chip and the optics that focus its light into the fiber. The ROSA, or Receive Optical Sub-Assembly, contains the photodiode that catches incoming light and the circuitry that turns it back into an electrical signal. Between them sits an electronic chip, principally a DSP (digital signal processor), which cleans up, times, and corrects the signal on both the way in and the way out.

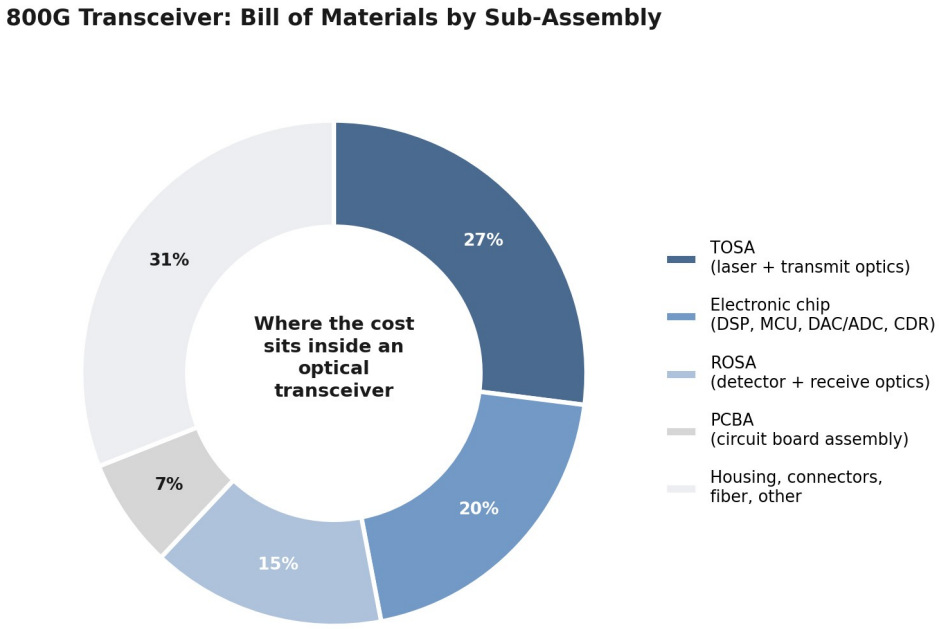

Figure 2. Cost composition of a representative 800G optical transceiver.
Source: Fiber Mall, Nomura Research, OsloMet Capital Group

The TOSA, where the laser lives, is the single largest cost line at roughly 27% of the bill of materials, with the electronic chip close behind at around 20% and the ROSA at 15%. But cost share understates the laser's real importance. The electronic chip and the housing can be sourced from any number of qualified suppliers. The laser cannot: without it, nothing else in the module functions, and it is by far the hardest single component in the transceiver to manufacture at the reliability level a hyperscale customer requires. That mismatch, modest cost share, outsized structural importance, is the reason laser supply specifically, not transceiver assembly generally, is where the sector's real scarcity sits.

SECTION 4

The Laser Is the Hard Part

Laser chips used in optical networking come in a handful of types, and the choice of type is really a trade-off between speed, distance, cost, and manufacturing difficulty. VCSELs are cheap and fine for very short reach. FP and DFB lasers extend that reach considerably. For the high-speed, longer-distance links an AI cluster actually needs, the industry has settled on two competing architectures, and the difference between them explains most of the competitive dynamics in the sector today.

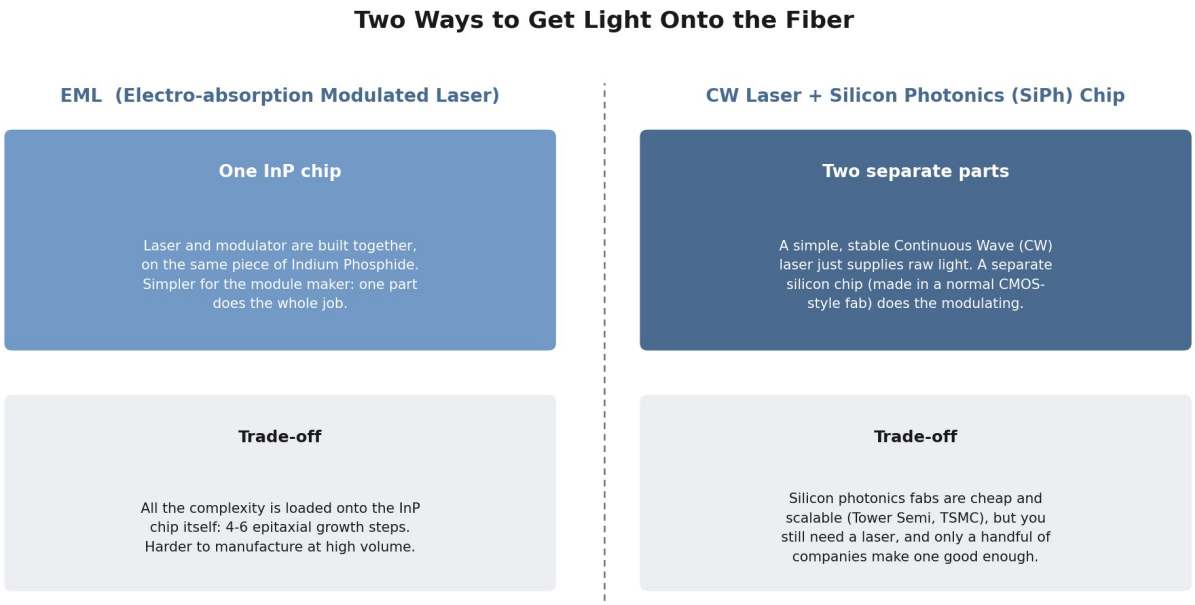

Figure 3. The two dominant architectures for getting a high-speed signal onto fiber.
Source: Nomura Research, company disclosures, OsloMet Capital Group

EML: everything on one chip

An EML, or Electro-absorption Modulated Laser, integrates the light source and the modulator that encodes data onto that light, monolithically, on a single piece of indium phosphide. It is currently the dominant, higher-margin solution for high-speed transceivers, and it is conceptually simple for the module maker: one part does the whole job. The cost of that simplicity is loaded entirely onto the chip itself. EML production typically requires four to six separate epitaxial growth steps and comparably demanding lithography, etching, and testing stages, which is why EML manufacturing capacity is hard to stand up quickly, and why it remains concentrated among a small number of established players.

CW plus silicon photonics: split the job in two

The alternative approach uses a simple, stable Continuous Wave (CW) laser that does nothing but supply steady light, paired with a separate silicon photonics (SiPh) chip that does the modulating. The appeal is manufacturing economics: silicon photonics chips can be built in ordinary, CMOS-style semiconductor foundries (Tower Semiconductor and TSMC are the leading examples), which are far cheaper to scale than a dedicated InP fab, because they reuse decades of mature silicon chip-making know-how. But silicon, on its own, cannot generate light. As one industry expert put it plainly in an interview conducted for this research: silicon photonics is a passive medium, and the missing piece in that equation is still a laser source. Splitting the architecture in two does not remove the laser bottleneck; it just isolates it more clearly, which is precisely why so much of the industry's competitive tension concentrates on a small number of laser suppliers, chiefly Lumentum and Coherent, rather than being spread evenly across the many companies capable of building the rest of a transceiver.

SECTION 5

Speed & Form Factor: The Two Axes of Change

The sector is moving along two separate axes at once, and conflating them is a common source of confusion. The first axis is raw speed: 400G, 800G, 1.6T, and now 3.2T, each roughly doubling the data rate of the last. The second axis is form factor: how tightly the optics are physically integrated with the switch chip they connect to, running from today's swappable pluggable module toward fully co-packaged optics.

Speed: units still skew legacy, even three years out

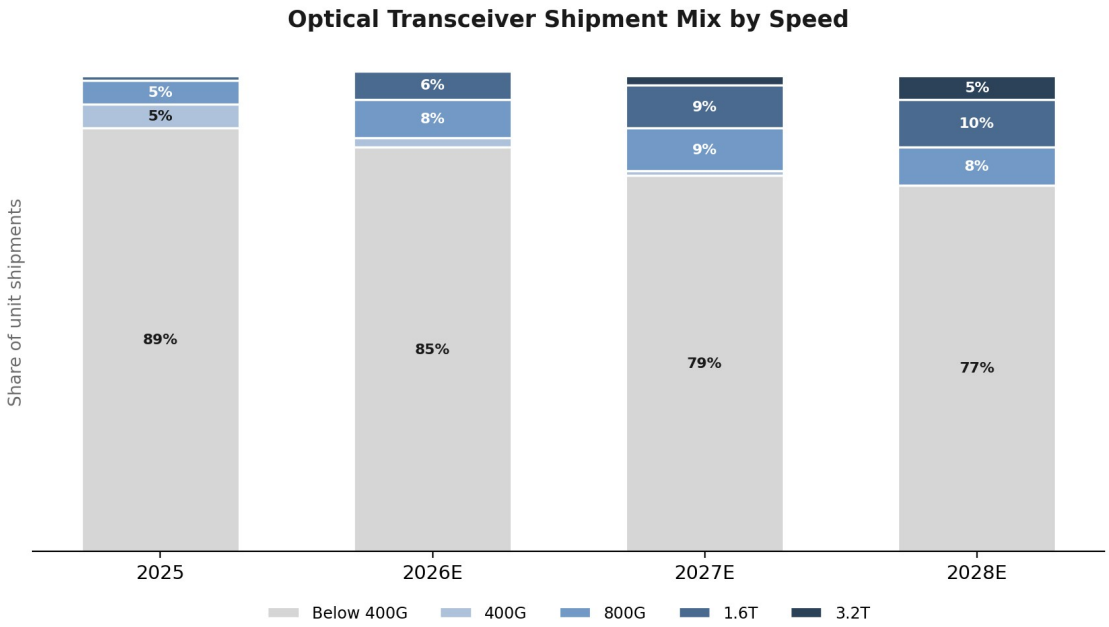

Figure 4. Modeled optical transceiver shipment mix by speed, 2025 to 2028E.
Source: Goldman Sachs Global Investment Research, OsloMet Capital Group

Even on 2028E modeled shipments, transceivers below 400G still account for the clear majority of units sold. The AI-specific story, the 800G, 1.6T, and emerging 3.2T tiers, is a value and margin story well before it is a unit-volume story: those higher speeds carry far higher price tags per unit even while remaining a minority of total shipments. This matters for how to read any given company's revenue growth. Strong dollar growth in a supplier's AI-facing segment can coexist with a business that is still, by unit count, dominated by legacy-speed product.

Form factor: where the signal-cleanup chip physically lives

Every AI server rack today mostly uses pluggable transceivers: self-contained modules a technician can swap in and out of a switch port, with the DSP and laser both housed inside. The industry has been experimenting with progressively tighter integration ever since, for one overriding reason: power. Every DSP and every connector consumes power and adds cost, and the industry-wide target price point (often cited as roughly $0.10 per gigabit) is not reachable with pluggables alone.

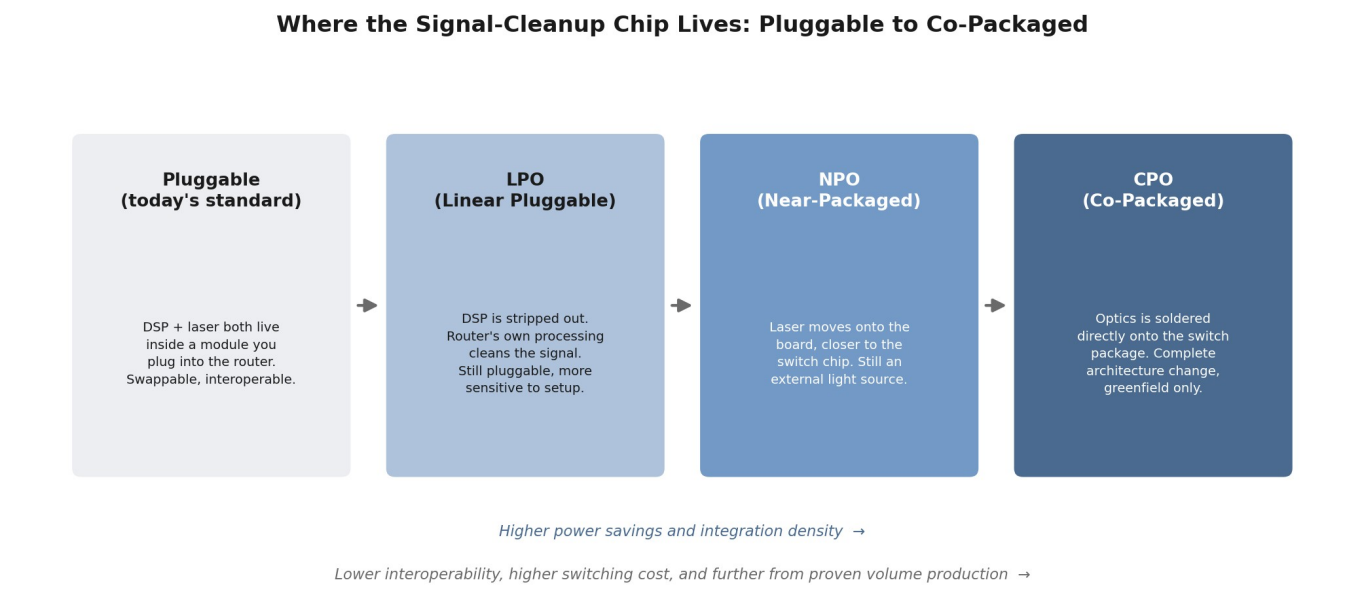

Figure 5. The progression from pluggable optics toward co-packaged optics (CPO).
Source: Expert interviews conducted for this research, OsloMet Capital Group

LPO strips the DSP out of the module and relies on the switch's own processing to clean up the signal instead, still pluggable, but more sensitive to the specific router it is paired with, and requiring close cooperation with that router's manufacturer. NPO moves the laser itself onto the board, physically closer to the switch chip, while still using an external light source. CPO is the most radical step: the optics are soldered directly onto the switch package, eliminating the pluggable form factor altogether. This is not an upgrade path. A switch built for pluggables cannot later be retrofitted for CPO; it requires a greenfield design from the start, which is the central reason CPO adoption, discussed further in Section 8, is a genuinely contested timeline rather than a settled one.

SECTION 6

The Upstream Squeeze: Why Indium Phosphide Is Scarce

Every EML laser, and every CW laser feeding a silicon photonics chip, starts life as a wafer of indium phosphide. Tracing that wafer back to its raw inputs reveals a chain that is short, but genuinely difficult to scale quickly, and the difficulty is not evenly distributed across it.

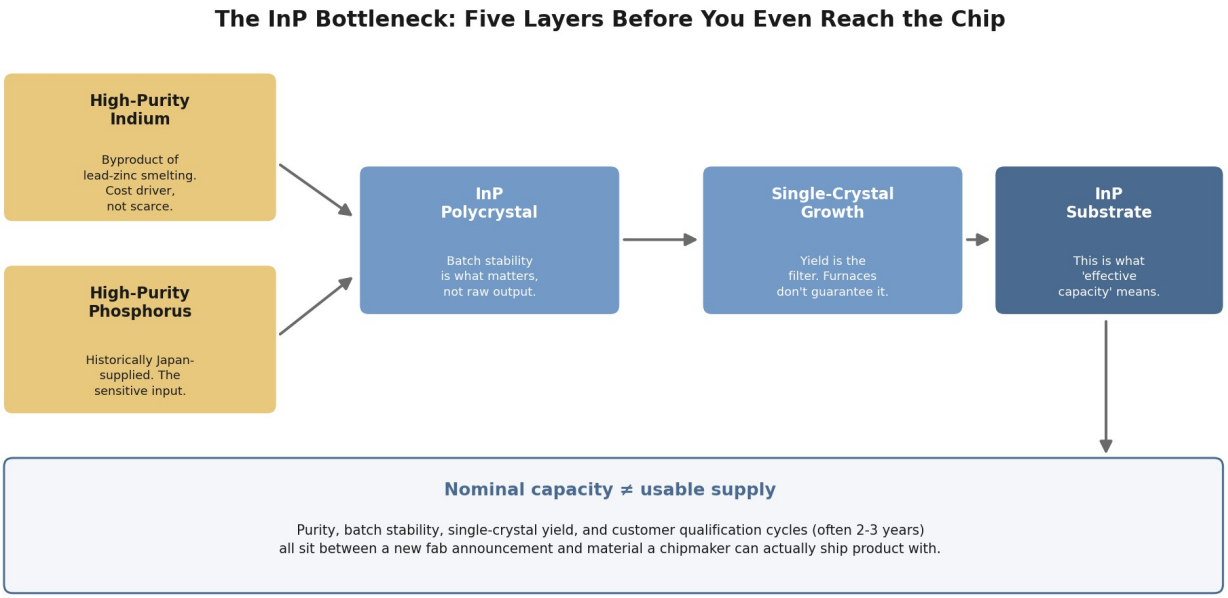

Figure 6. The indium phosphide supply chain, from raw material to usable substrate.
Source: Global Semi Research, industry disclosures, OsloMet Capital Group

Indium itself, despite the attention it gets, is more of a pricing variable than a hard supply constraint. It is produced mainly as a byproduct of lead-zinc smelting, and China is a major supplier; between primary and recycled indium, there is generally enough of it to go around, even though its price has risen sharply. Phosphorus is the more sensitive input, historically dependent on a small number of Japanese suppliers of the high-purity grade InP production requires.

But the real constraint sits further downstream in this same chain, at the polycrystal and single-crystal growth stages. Turning high-purity raw material into a usable substrate is not primarily an equipment problem: furnaces can be bought and fabs can be built. It is a process-know-how problem, involving temperature control, batch-to-batch stability, and single-crystal yield, and that know-how is accumulated over years, not purchased. This is why announced capacity expansions across the industry should be read carefully: nominal capacity is not the same thing as usable, qualified supply, and the gap between the two is often where a bottleneck actually persists even as headline numbers suggest the shortage is easing.

A separate, related constraint sits on the customer side rather than the production side: qualification. Before a major customer will use a new laser or optical component supplier in

volume, that supplier typically must pass a reliability testing cycle, referencing Telcordia and MSA (multi-source agreement) industry standards, that one industry expert consulted for this research put at around three years. This is the main reason raw manufacturing capacity, wherever it exists, does not translate quickly into market share: a producer can have functioning output and still be multiple years away from shipping into a qualified, revenue-generating supply chain.

SECTION 7

Sizing the Market: Growth and Price Decay, Together

The optical module market is growing quickly by almost any measure, but the dollar-value growth curve is not a straight line, and understanding why requires looking at unit growth and pricing together rather than separately.

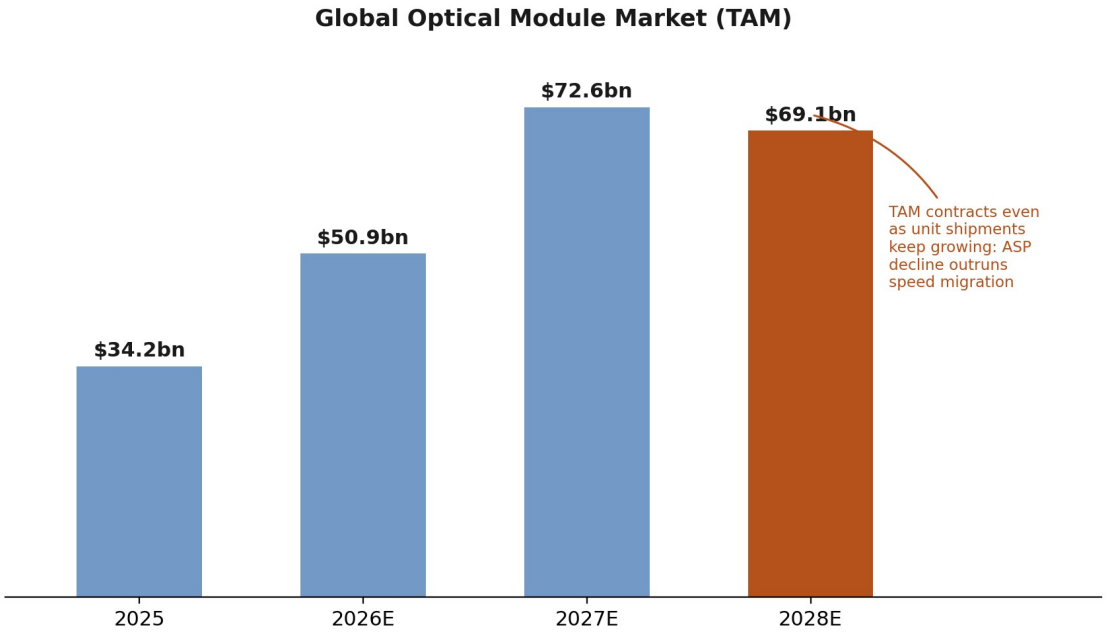

Figure 7. Global optical module total addressable market, 2025 to 2028E.
Source: Goldman Sachs Global Investment Research, OsloMet Capital Group

Modeled industry revenue grows from roughly $34.2bn in 2025 to a peak of around $72.6bn in 2027E, before contracting to approximately $69.1bn in 2028E, even as unit shipments keep rising throughout the period. The explanation sits entirely in pricing.

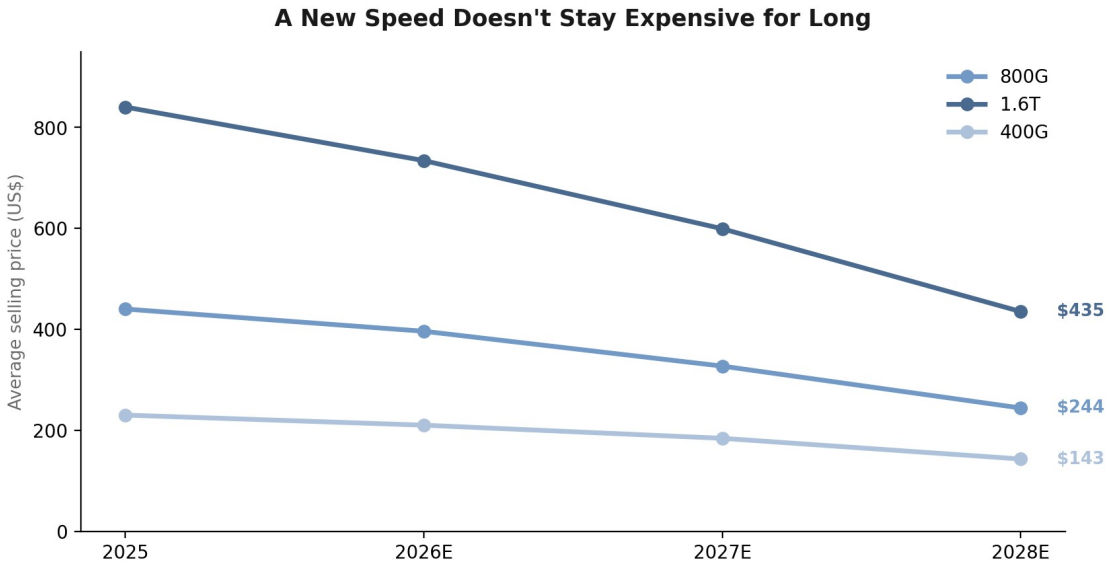

Figure 8. Average selling price by speed tier, 2025 to 2028E.
Source: Goldman Sachs Global Investment Research, OsloMet Capital Group

Every new speed tier follows the same arc: a high price at launch, followed by rapid, sustained price decay as production scales, competition intensifies, and the technology matures. On these modeled figures, 800G pricing falls by roughly a quarter between 2025 and 2028E, and 1.6T pricing, despite still being a comparatively new product, falls by roughly a fifth in its fourth full year alone. This is the ordinary pattern of the optical component cycle repeating itself at each new speed generation, and it is the main reason unit growth and dollar growth can, and do, diverge.

The laser chip layer specifically, further upstream and smaller in absolute size, is estimated to be growing faster than the finished-module market around it. Industry estimates put the global laser chip market at roughly $2.6bn in 2024, rising toward approximately $22.9bn by 2030E, a compound growth rate above 40% a year. Within that, EML and CW lasers together are estimated to grow even faster, from roughly $970mn in 2024 to close to $20.8bn by 2030E, rising from around 38% of the total laser chip market today to more than 90% of it by the end of the decade. This is the layer of the value chain where structural scarcity, not just cyclical demand, does most of the explaining.

SECTION 8

Where Informed Opinion Actually Disagrees

Most of what appears in sell-side research on this sector agrees on the big picture: AI demand is real, supply is tight, and the leading names are gaining share. Independent industry experts consulted for this research, people with direct technical and strategic experience inside the sector rather than a financial stake in any single stock, agreed with the demand story but pushed back, specifically and independently of one another, on three points.

1. Where the real bottleneck sits

Most bank research frames the binding constraint as raw wafer and chip manufacturing capacity, the argument behind the current wave of capital spending across the sector's leading names. Two independent experts consulted separately for this research instead pointed to packaging, testing, and customer qualification cycles as the actual chokepoint, a different point in the chain that new fab capacity does not directly resolve. If that view is correct, capital currently being spent on wafer and chip capacity may not relieve the constraint on the timeline the market is currently pricing in.

2. The timing of CPO adoption

Company management commentary and most sell-side modeling point to meaningful co-packaged optics revenue arriving in 2027 and scaling through 2028. Two independent experts, including a former senior technology strategy director at Lumentum, separately estimated that broad, high-volume hyperscaler adoption is realistically five or more years out from mid-2026, closer to 2030 or later, with one citing a recently published third-party industry report (LightCounting's Cloud Data Center Optics report) projecting that classic pluggable transceivers still hold more than half the market through 2031. A meaningful share of the revenue currently modeled into 2027 and 2028 for CPO-focused product lines rests on a timeline that people close to the technology, with no financial stake in the outcome, do not think is realistic.

3. A structural concentration risk nobody prices explicitly

One expert described the small number of companies currently capable of producing qualified, transceiver-grade lasers, effectively two or three firms, as functioning close to a duopoly that, in their words, has to stay alive for the broader AI infrastructure buildout to proceed on its current roadmap. This is a systemic dependency, distinct from ordinary competitive risk between two rivals, and it does not appear as a named risk factor in any of the sell-side research reviewed for this brief.

CLOSING

Conclusion: What to Watch Next

None of the above is a reason to dismiss the bull case for the sector; the physical demand driver in Section 1 is real and is not going away. It is a reason to hold the consensus view and the informed dissent from it side by side, and to track the specific, checkable signals that would resolve the disagreement one way or the other:

- **Qualification announcements:** watch for named suppliers publicly confirmed to have passed Telcordia and MSA reliability testing with a major customer, the clearest sign that qualification friction is easing faster or slower than expected.
- **6-inch InP yield data:** track disclosed yield and output data specifically, not just capacity announcements, since the gap between the two is where the real bottleneck has consistently been shown to sit.
- **First non-Nvidia CPO shipments at volume:** the clearest possible falsification of the adoption-timeline debate in Section 8 is any hyperscaler beyond Nvidia's own compute nodes reporting real, high-volume CPO shipments.
- **LightCounting's own published Cloud Data Center Optics report:** this third-party report was cited independently in expert conversation as projecting pluggables above 50% share through 2031; worth reading directly rather than through a single secondhand summary.
- **Quarterly ASP prints:** watch whether each quarter's actual pricing tracks, beats, or undershoots the decay curve in Figure 8; a slower-than-modeled decline would be the clearest sign that shortage conditions are more durable than the consensus base case assumes.
- **US import-restriction policy news flow:** any US restriction on Chinese-made optical modules would reshape competitive positioning across the entire chain discussed in Section 2, independent of anything happening on the technology side.

Glossary

InP	Indium phosphide. The semiconductor material almost all high-speed laser and detector chips are built on, in place of the silicon used for ordinary logic chips.
LD / PD	Laser diode (the light source, transmit side) and photodiode (the light detector, receive side).
EML	Electro-absorption Modulated Laser. A laser and its modulator built together on one InP chip. Today's high-speed workhorse.
CW laser	Continuous Wave laser. A simple, always-on light source with no modulator built in; used alongside a separate silicon photonics chip.
SiPh	Silicon photonics. Using standard, CMOS-style silicon chip manufacturing to build the modulator and routing around a laser, instead of building everything on InP.
TOSA / ROSA	Transmit / Receive Optical Sub-Assembly. The two halves of a transceiver: TOSA houses the laser, ROSA houses the detector.
DSP	Digital Signal Processor. Cleans up and times the electrical signal before and after it becomes light. Sits inside a pluggable module, or gets moved elsewhere in LPO/CPO designs.
Pluggable / LPO / NPO / CPO	The four stages of how tightly the optics are integrated with the switch chip, from a swappable plug-in module (pluggable) to optics soldered directly onto the switch package (CPO). See Section 5.
ASP	Average Selling Price. The per-unit price a transceiver sells for; the key variable behind why market value and unit volume can move in different directions.
TAM	Total Addressable Market. The full dollar size of a market, here referring to global optical module or laser chip revenue.
Scale-up / scale-out / scale-across	The three distances optical connectivity now has to cover: GPU-to-GPU inside a server (scale-up), rack-to-rack inside a data center (scale-out), and data-center-to-data-center (scale-across).
Qualification (Telcordia / MSA)	The reliability testing cycle a new optical component or supplier must pass before a major customer will use it in volume. Typically cited at around three years for transceiver-grade lasers.

Sources

AlphaSense. (2026, August 14). *Former director of technology strategy at Lumentum believes photonics will play a pivotal role in AI and telecom advancements* [Expert interview transcript]. Tegus Expert Insights.

Arya, V., Jang, D., Mani, M., & Pharr, L. (2026, August 11). *Lumentum Holdings: First take: Solid growth, execution, but moderating leverage, fair valuation* [Equity research report]. BofA Securities.

Chang, A., Jeng, V., & Song, T. (2026, August 11). *Optical networking: Top 3 investor debates on impact of potential US import restrictions on Chinese optical transceiver modules* [Equity research report]. Goldman Sachs.

Genovese, M. (2026a, August 7). *Applied Optoelectronics, Inc. (AAOI): 2Q26 shows milestones confirming bullish thesis; capacity constraints in focus but story on track* [Equity research report]. Rosenblatt Securities.

Genovese, M. (2026b, August 13). *Coherent Corp. (COHR): Strong quarter & guide; nearly all the benefit of the optical industry revolution still ahead* [Equity research report]. Rosenblatt Securities.

Genovese, M. (2026c, August 11). *Lumentum Holdings, Inc. (LITE): Demand keeps growing; LITE has best execution, good answers to CPO & pricing questions* [Equity research report]. Rosenblatt Securities.

Global Semi Research. (2026). *One layer further upstream of optical chips*. Substack. *(Author identity not disclosed beyond this byline; treat with the reduced confidence noted when this source was first reviewed.)*

Ng, M., Goenka, Y., & Shema, L. (2026, July 28). *Celestica Inc. (CLS): 2Q26 review: Higher 2027E estimates on OpenAI, AMD programs and supply availability* [Equity research report]. Goldman Sachs.
Nomura Global Markets Research. (2026, July 8). *Yuanjie Semiconductor Technology* [Equity research report]. Nomura.

P Equity Research. (2026, August 16). *Expert call #2: Optical networking* [Expert interview transcript]. *(Interviewee identity withheld under NDA; interview conducted August 13, 2026.)*

Sylla, P., & Malik, A. (2026, August 11). *Lumentum Holdings Inc (LITE.O): Management callback notes* [Equity research flash note]. Citi Research.