Independent market research and competitive analysis of next-generation business and technology solutions for service providers and vendors

HEAVY READING

WHITE PAPER

Nearing the Shannon Limit: Terabit Coherent Optical Networks

A Heavy Reading white paper produced for Nokia

NOKIA

AUTHOR: STERLING PERRIN, SENIOR PRINCIPAL ANALYST, HEAVY READING

INTRODUCTION

As commercial optical systems approach the capacity and spectral efficiency thresholds imposed by the Shannon limit, service providers increasingly must seek performance gains in new and innovative ways. This is particularly true of long-haul and subsea networks, where top optical system requirements include possessing the longest reach, highest system capacity, and highest spectral efficiency over fiber. These networks typically see the earliest adoption of the highest performance optical systems.

This white paper takes an in-depth look at the current state and future trajectory of coherent optics. The paper describes the bifurcation of optics along two tracks: high performance optics and low power pluggable optics. Outlining the practical implications of the Shannon limit, Heavy Reading then discusses those implications in the context of high performance optics. After discussing key enabling innovations in high performance, this paper concludes with a look at what comes next.

TWO DISTINCT COHERENT OPTICAL MARKETS

Coherent detection was the springboard that launched the DWDM capacity revolution of the last decade. Starting in 2010 with 32 Gbaud digital signal processors (DSPs) and using QPSK modulation, systems were able to get 8Tbps of capacity on a fiber pair using the C-band spectrum. From here, advances in modulation formats combined with higher baud rates led to higher and higher capacities on a fiber.

Higher order modulation boosted channel capacity from 100 to 200Gbps in 2013. Then, channel rates made a big jump to 400Gbps starting in 2015, aided by higher order modulation combined with an increase in the baud rate. 600G was introduced in 2017 and combined 64–67 Gbaud symbol rates with 64QAM modulation. State-of-the-art commercial technology today consists of 7nm DSPs operating at 90–96 Gbaud (depending on the supplier). These systems are capable of operating up to 800Gbps per wavelength capacity when using 64QAM modulation.

Two tracks of coherent DSP development have emerged

Significantly, on the path from 100G to 800G, the coherent DWDM bifurcated along two distinct paths, with one set of innovations aimed at the high performance market and a separate set of innovations aimed at an emerging low power and pluggable-focused market. This bifurcation started around 2015, beginning at the 200Gbps data rate.

High performance market

High performance requirements include the following:

- Longest possible reach for any given capacity
- Highest achievable system capacity
- Greatest spectral efficiency per fiber
- Lowest system and network cost that meets required performance

High performance was the initial market for coherent detection, as chromatic dispersion (CD) and polarization mode dispersion (PMD) made long-haul transmission (or even metro transmission) impossible at 100Gbps using direct detect optics. The first coherent 100Gbps systems, by contrast, matched typical 10G performance in terms of distance, achieving 1,500+km reaches in real-world deployments.

Coherent systems not only increased per-channel capacity by 10x but also boosted spectral efficiency, measured as bit per second per hertz of spectrum. Compared to direct detect 10G, coherent 100Gbps using dual polarization increased spectral efficiency from 0.2 bps/Hz to 2 bps/Hz—also a 10x improvement.

These high performance requirements have continued in coherent systems in subsequent technology generations, particularly in long-haul and ultra-long-haul terrestrial and subsea applications, in which distance requirements range from thousands of kilometers to as high as 10,000+km for some subsea routes.

Lowest cost is always an important requirement, but for the high performance segments, cost must be considered in the context of the other primary factors. For one, if a system cannot reliably connect two endpoints due to limited reach, then its cost is irrelevant. Additionally, a higher cost system may still yield the lowest cost network build if, for example, longer system reach reduces regeneration sites along a route or higher spectral efficiency eliminates the need for a network overbuild.

In high performance systems, space, power, and interoperability considerations are secondary to the key requirements described above. Historically, high performance coherent technologies are introduced in embedded modules that use vendor-proprietary forward error correction (FEC) that delivers the greatest performance for that vendor's specific system.

Lower power pluggable market

Low power and pluggable-focused market requirements include the following:

- Lowest cost per bit for any given data rate
- Lowest power consumption
- Smallest footprint
- Modularity/pluggability
- Degrees of standardization

Initial coherent systems were "one size fits all," regardless of application. But the introduction of high order modulation formats such as 16QAM and 64QAM, increases in the baud rate, and miniaturization aided by silicon photonics made way for purpose-built transponders and DWDM systems geared specifically for data center interconnect and metro applications.

Coherent module development began diverging at 200Gbps with the introduction of higher performance embedded and lower power pluggable options. However, the 400G data rate marks a major milestone in coherent bifurcation driven particularly by the Optical Internetworking Forum's (OIF's) 400ZR implementation agreement, published in April 2020. 400ZR specifies pluggable coherent 400 Gigabit Ethernet modules that can be housed in QSFP-DD and OSFP form factors. Lower power limits maximum reach to 120km with amplification, which is in line with the targeted metro DCI applications. OIF standardization enables interoperability and opens the door for massive volume deployments.

Other 400G open coherent pluggable groups and projects include the Open ROADM Multi-Source Agreement (MSA) and the Open ZR+ MSA. OIF has kicked off the 800ZR project to define coherent 800Gbps. Additionally, some vendors are building 400G coherent pluggables with proprietary FEC for extended reach into regional/long haul.

Figure 1 outlines major open coherent pluggable optics module options and key specifications. In addition, some suppliers may also sell proprietary modules, primarily in the CFP2 form factor.

Figure 1: Open pluggable coherent module options

	400ZR	Open ZR+	Open ROADM	800ZR
Target application	Edge DCI	Metro DCI and carrier	Carrier metro/ regional	Edge DCI
Client traffic	400GE	100–400GE	100–400GE and OTN	800GE
Target reach @ 400Gbps	120km	Metro/regional	500km	80+km
Form factor	QSFP-DD/OSFP	QSFP-DD/OSFP or CFP2	CFP2 or other	QSFP-DD/ OSFP
Soft-decision FEC (SD-FEC)	Concatenated FEC (CFEC)	openFEC (oFEC)	oFEC	
Standards/ MSA	OIF	OpenZR+ MSA	Open ROADM MSA, ITU-T	OIF Project

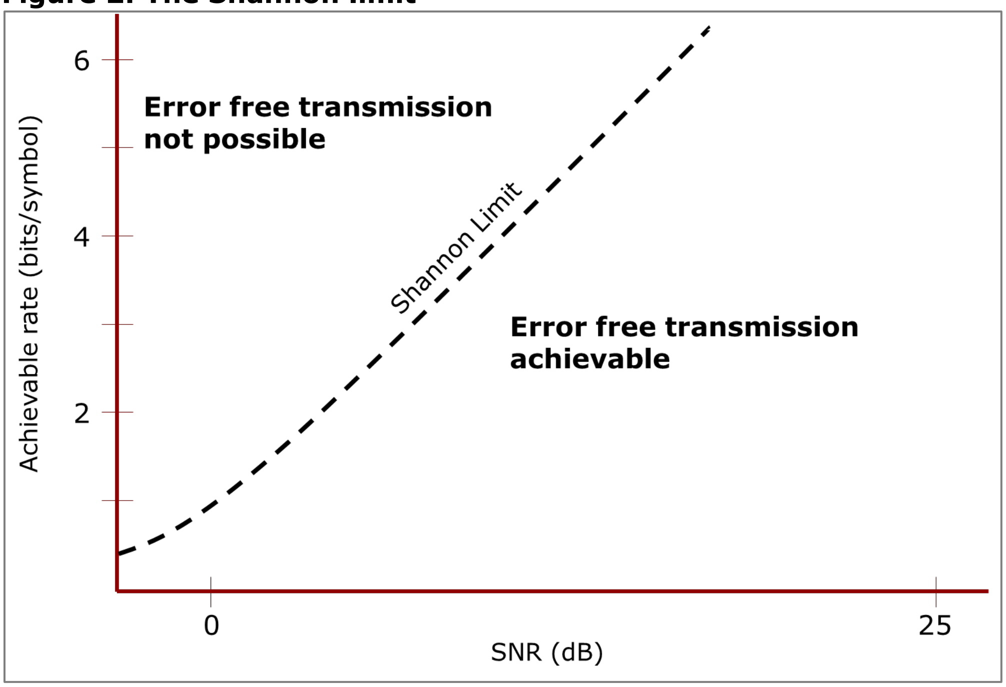
Source: Heavy Reading, 2022

PRACTICAL IMPLICATIONS OF THE SHANNON LIMIT

Thanks to coherent detection, wavelength capacity has been increasing steadily over the past decade. The CAGR for coherent WDM channel capacity is 19% from the introduction of coherent 100G to the present (2010–22). Vendor announcements and plans indicate that WDM channel capacity will continue to increase, with 1.2Tbps as the next channel rate expected to be commercialized beginning in 2023.

However, capacity cannot be expanded infinitely—even in theory. Claude Shannon's 1948 paper defined the theoretical maximum data rate over a communications medium in the presence of noise. Regardless of the modulation format used, all transmission has a limit beyond which information capacity cannot increase. This threshold that caps theoretical capacity based on the signal-noise ratio (SNR) is called the Shannon limit (see **Figure 2**).

Figure 2: The Shannon limit



Source: Heavy Reading, 2022

Plotting the evolution of data rates, baud rates, and modulation rates over the past decade illustrates the practical limitations imposed by Shannon's physics. As detailed in **Figure 3**, advances in higher order modulation formats (e.g., from QPSK to 16QAM to 64QAM) drove C-band system capacities higher from 2010 to 2016, with system capacity increasing at 22% for the period (essentially, at the same rate as wavelength capacity). Spectral efficiency also increased during this time, rising from 2 bps/Hz to 8 bps/Hz.

Figure 3: Coherent DWDM capacity evolution

Max data rate	Baud rate	Max C-band cap (4.8THz)	Max channels	WDM channel spacing	Max spec efficiency	Introduction
100Gbps	32 Gbaud	9.6Tbps	96	50GHz	2	2010–11
200Gbps	32 Gbaud	19.2Tbps	96	50GHz	4	2013–14
400Gbps	32–45 Gbaud	30.4Tbps	76	62.5GHz	6.3	2015–16
600Gbps	64+ Gbaud	38.4Tbps	64	75GHz	8	2017–18
800Gbps	90-100 Gbaud	33.6Tbps	48	100–125GHz	6–7	2019–22
1.2Tbps	130Gbaud	38.4Tbps	32	150GHz	8	2023 expected

Source: Heavy Reading, 2022

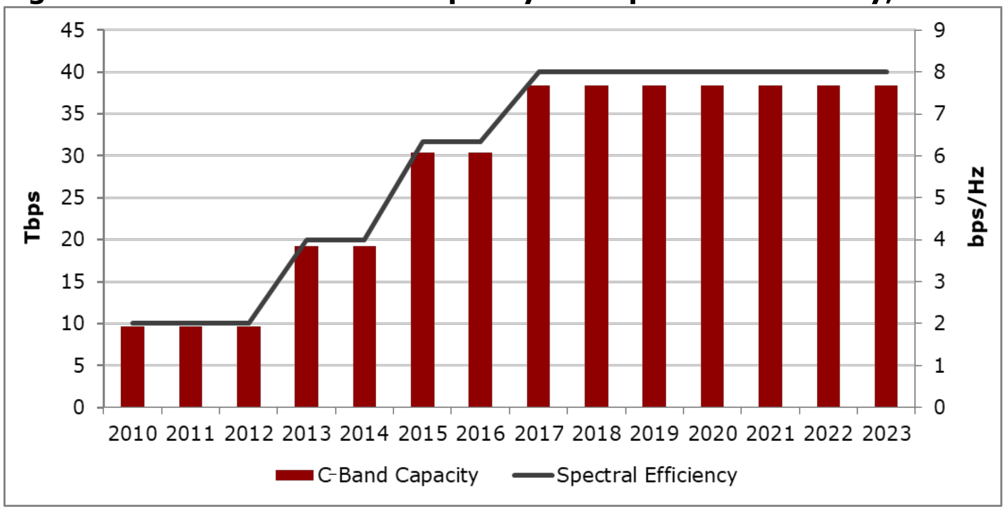
However, higher order modulations impose significant penalties in terms of reach due to the sharp increase in sensitivity to noise and nonlinearities as modulation complexity increases. Doubling channel capacity by moving from QPSK to 16QAM, for example, results in a reach reduction of approximately 75%. Beyond 64QAM, reach is too restrictive for wide area applications.

As a result, to increase channel capacity, DWDM suppliers must rely on increasing the baud rate (also called the symbol rate). The key benefit of baud rate versus modulation format is that increasing the baud rate boosts achievable reach for any given data rate compared to the previous generation. The baud rate, however, imposes a different cost. Higher baud rates deliver higher headline capacities but require wider channel widths and thus are typically unable to increase spectral efficiency.

This significant capacity limitation has been evident in DWDM commercialization over the past five years. Since the introduction of 64–67 Gbaud DSPs at 600Gbps, spectral efficiency has plateaued at 8 bps/Hz. This means that total C-band capacity has also been capped at 38.4Tbps, even as 90+ Gbaud DSPs have increased the headline channel capacity to 800Gbps.

The spectral efficiency cap also applies to proposed 5nm, 1.2Tbps DSPs that are expected to be introduced beginning in 2023 and for any future DSP introductions moving forward, regardless of maximum channel rate (e.g., 1.6Tbps, 2Tbps, and beyond). Plotting total C-band capacity and spectral efficiency through 2023 (projected), **Figure 4** illustrates the communications industry's capacity challenge.

Figure 4: Maximum C-band capacity and spectral efficiency, 2010–23



Source: Heavy Reading, 2021

HIGH PERFORMANCE MARKET IMPLICATIONS

The Shannon limit applies to all optical data rates and applications, but its effects are most immediate in the high performance market because of the priority placed on longest reach, highest system capacity, and greatest spectral efficiency. This section looks at these requirements in light of the Shannon limit.

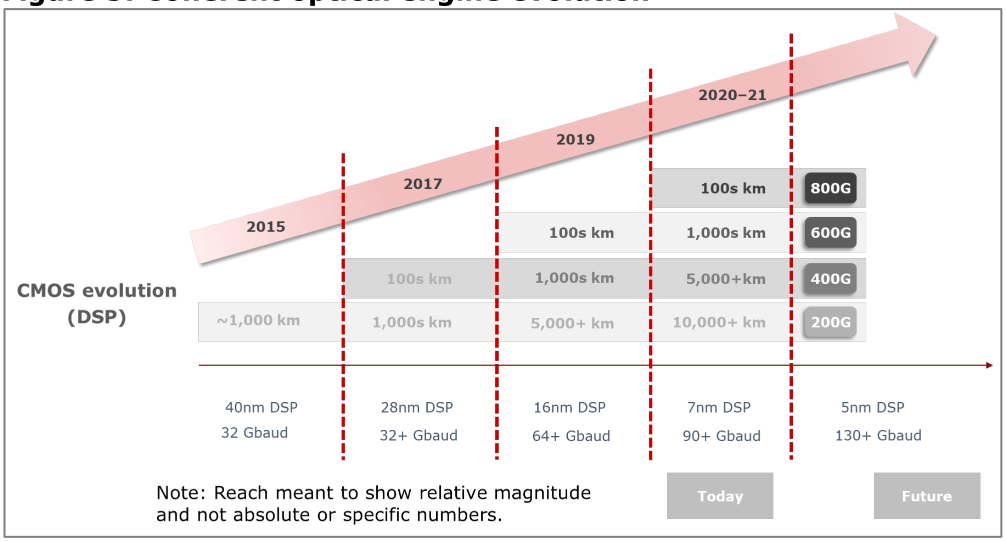
Reach

The top or "headline" data rate for a coherent generation can be misleading because this data rate is only achievable over relatively short distances that are insufficient for long-haul networks (or, typically, even regional networks). This does not mean, however, that the coherent technology evolution lacks relevance. By employing the newest baud rate but dialing down to lower order modulation formats, each new coherent generation provides greater reach than the previous generation for any data rate.

Figure 5 traces the evolution of coherent optical engines and attaches general distance metrics to the major data rates. Looking at the figure, systems based on 16nm/64–67 Gbaud DSPs achieve long-haul reaches at 400Gbps but must drop to metro distances at 600Gbps. Moving to 7nm/90+ Gbaud DSPs, coherent systems comfortably address long-haul reaches at 600Gbps (e.g., typically 1,500km). When dialing down to 400Gbps, these systems can achieve distances of thousands of kilometers.

Therefore, many long-haul and subsea service providers will be justified in migrating to the newest generation of coherent technology even though the headline data rate is not achievable (e.g., 800Gbps for today's generation). This migration trend will continue at least until spectral efficiency gains for long-haul/subsea networks have been exhausted.

Figure 5: Coherent optical engine evolution



Source: Heavy Reading, 2021

A key focus for Nokia has been to demonstrate the real-world performance of its Photonic Service Engine (PSE-V) super-coherent optics in field trials over live networks. Performed in collaboration with various network operators, these have typically been conducted in complex networks with multiple ROADM pass-throughs, over existing fiber plant such as single-mode optical fiber (SMF) or mixed fiber types, and with spectrally efficient 100GHz WDM channels.

In a trial with Orange, Nokia demonstrated PSE-V's operation at 600Gbps over a 914km network between Paris and Biarritz, across 13 spans and through multiple cascaded ROADMs. In the GlobalConnect trial, Nokia demonstrated transmission over a 781km ring across 11 fiber spans and through six contentionless-flex grid (CDC-F) ROADM nodes.

System capacity/spectral efficiency

As noted throughout this paper, beginning at 64–67 Gbaud, total achievable capacity in the C-band has plateaued at 38.4Tbps. Moving forward, service providers must seek new ways to increase optical network capacity. Significantly, this is not a theoretical discussion; rather, it is an issue that affects many service providers in high performance segments today.

One current solution is to add L-band transmission to the existing C-band. This C+L-band combination essentially doubles fiber capacity (~76.8Tbps) but comes at the expense of a new and parallel amplifier system supporting the new spectrum band, which operates at 1565–1625nm, just above the C-band. C+L-band transmission was first commercialized in subsea networks by hyperscalers, including Google and Facebook. Today, C+L-band is used in subsea as well as in terrestrial long-haul networks, including by North American cable multiple system operators (MSOs) Comcast, Charter, and Cox.

Another solution is to add more fibers—which is the most basic form of spatial division multiplexing (SDM). In planning new capacity, service providers will compare the feasibility and total costs of C+L-band versus laying fibers. For metro and edge networks, the cost analysis overwhelmingly favors adding fibers, and this fiber deployment strategy is certainly playing out as communications service providers (CSPs) plan capacity for 5G. However, in long-haul terrestrial and in subsea networks, fiber project costs are much higher (if feasible at all). For this reason, the C+L-band capacity option is specific to long-haul and subsea networks.

Beyond these two options, other possibilities have been discussed for years but are not close to commercialization. Adding the S-band (1460–1530nm range) would give service providers another block of spectrum for capacity, but costs and complexity challenges make S-band adoption unlikely for at least the medium term.

Similarly, SDM using multi-core fibers would boost fiber capacity by the number of fiber cores, but after many years, this too remains in the academic and research domains without a clear path to commercial adoption.

HIGH PERFORMANCE SYSTEM DIFFERENTIATORS

In high performance systems, service providers favor performance metrics over interoperability and standardization, and proprietary innovation will continue to trump standardization in this market for the foreseeable future. This section details several of the key differentiating technologies in high performance coherent optics.

90+ Gbaud DSP

The state-of-the-art in commercial coherent DWDM systems today includes 7nm DSP ASICs operating from 90 to 96 Gbaud (depending on vendor). Systems based on these DSPs are primarily targeted for use in regional and long-haul networks where maximum performance is needed. In long-haul links, these systems can operate at a per-channel data rate of 600Gbps in links up to approximately 1,500km, and they have a significantly greater reach of 3,000+km at a per-channel data rate of 400Gbps.

DSP ASIC designs take roughly 18–24 months, cost in the tens of millions of dollars, and skill sets are rare. The trend toward vertical integration (detailed below) adds further complexity. As of 1Q22, just three suppliers—Ciena, Infinera, and Nokia—have commercial DSPs in the 7nm generation. (Cisco has announced plans for a next-generation 5nm DSP, but it is not available.) Thus, availability of the DSP alone is a major market differentiator.

Beyond availability, there is further differentiation within this limited supplier base. Nokia, for example, operates at a slightly lower baud rate (90 Gbaud) that uses a 100GHz channel width that is compatible with existing ROADM networks built for 50GHz spacing. It allows more channels per fiber in the C-band compared to 95–96 Gbaud solutions that require WDM channel spacings of 112.5–125GHz. Channel compatibility will be an important consideration for some operators. Additionally, Nokia's DSP is targeted at high performance regional and long-haul applications where operation is limited to 600Gbps or less and thus does not support the 64QAM 800Gbps operation achievable in metro/regional applications.

Nokia's approach considers the availability of 400G pluggable coherent optics built for metro applications that are more modular and typically lower cost. According to Nokia, eliminating the 64QAM modulation function results in a 15% reduction in size and a 20% reduction in power consumption. Space/power reduction will be an important consideration for some service providers, but others may require the full 800Gbps data rate for their metro/regional applications.

Vertical integration with co-packaging

Vertical integration has become a hallmark of the high performance market as coherent optics suppliers have consolidated DSP design, photonic integration, and DSP/frontend co-packaging expertise in-house. The initial driver for vertical integration was the move from analog coherent optics (ACO) designs, in which the DSP is located outside the module, to digital coherent optics (DCO) designs, in which the DSP is located within the optical module itself. DCO design is required for pluggable optics, and this has been a key enabler for 400G pluggable optics, including 400ZR. However, pluggability is not the only driver. Owning both the DSP and the optics allows tight coordination throughout the R&D process that yields the highest performance systems.

This trend is especially true at today's 90 Gbaud coherent generation and any generation beyond, where tight integration and close connectivity between components also require co-packaging expertise. Moving forward, each new coherent generation will be DCO due to these stringent technical requirements.

The high cost of DSP ASIC design combined with the requirement for vertical integration has resulted in extreme barriers to entry in the high performance market. While the cost-focused pluggable optics market is experiencing growth in suppliers, the number of viable suppliers in the high performance market has declined.

All three suppliers (Ciena, Infinera, and Nokia) that have generally available DWDM systems at 90+ Gbaud have adopted a vertically integrated model.

Probabilistic constellation shaping

Probabilistic constellation shaping (PCS) is one technique increasingly applied to modulation symbols to boost optical performance and close the gap between what is theoretically defined by the Shannon limit and what exists in actual transmission systems. Standard modulation formats (e.g., 16QAM, 64QAM) use all points in the QAM constellation map with equal probability, producing a "flat" or "square" distribution. The problem is that Shannon's theorem is based not on equal distribution, but rather on a Gaussian noise model in which noise is not uniform. Significantly, the Gaussian model, not even distribution, is the optimal model for maximizing performance.

With PCS, complex algorithms assign different probabilities to different constellation points when mapping the bits to symbols, with higher probabilities given to constellation points with lower powers. The result is a distribution of constellation points that closely matches the optimal Gaussian distribution. According to Nokia—which was the first vendor to adopt PCS in optical networks in the 67 Gbaud generation of DSPs—PCS implementation improves reach by 1dB, or about 25%. This is a significant improvement in the context of networks that seek to squeeze the highest capacity and greatest reach possible.

FEC

FEC algorithms have a long history in DWDM networks. Although they have historically been proprietary, standardized FEC algorithms are being adopted in the cost-focused coherent pluggables market, most notably in 400ZR and Open ROADM. These standardized algorithms provide "good enough" performance, and standardized FEC is essential for vendor interoperability.

However, in high performance networks, service providers seek the best performance possible, and vendor interoperability is not a requirement. Here, FEC is another feature that will differ by supplier and will remain vendor-proprietary for the foreseeable future. One key area of differentiation is maximum error-free transmission distance. Another differentiating factor is the number of overhead bits that FEC consumes.

Historically, high gain FEC used in high performance coherent applications requires more than 20% overhead—accounting for the bulk of the difference between the coherent bit rate and the actual data rate. FEC innovations that lower overhead can be co-optimized with PCS to either allow more payload capacity to be transmitted for a given reach or increase reach for a given payload speed. In both cases, this can help service providers squeeze greater performance from long-haul networks.

WHAT'S NEXT IN COHERENT OPTICS?

With the introduction of 7nm DSPs, the industry is looking to what's next for the future of coherent optics. Significant industry debate exists around whether the next step is to achieve 5nm/130 Gbaud or to skip 5nm and design for 3nm/180 Gbaud. Driving this debate is the reality that coherent optics already operate close to the Shannon limit, and each new investment cycle moving forward will result in smaller performance gains. The C-band is full, and today's 7nm/90+ Gbaud is capable of long-haul distances at 600Gbps.

Cisco has already announced plans for a 5nm generation, but the rest of the industry remains undecided. Timing of 3nm availability will be a major factor in these decisions. While the technology ecosystem for 5nm coherent DSP is mature today, it remains to be seen whether the same can be said for key technologies needed for 3nm coherent, including analog-to-digital and digital-to-analog converters (ADCs and DACs), fab maturity for mixed-signal ASICs, and the complementary high speed optics needed for 180–200 Gbaud operation.

Taiwan Semiconductor Manufacturing Company (TSMC) and Samsung are the main complementary metal-oxide-semiconductor (CMOS) manufacturers, with TSMC stating plans to start 3nm production in 2H22. However, technical challenges and the complex manufacturing process for building 3nm CMOS (as noted above) will likely cause delays for both manufacturers, pushing initial availability to 2023.

Additionally, demand is driven by digital ICs for mobile devices with massive volumes. Optical networking, with lower volumes, is a lower priority and is at least one year behind any initial schedule.

CONCLUSIONS

As the coherent optical market has matured over the past decade, it has diverged along two distinct development tracks: one aimed at high performance optics primarily for long-haul/ultra-long-haul terrestrial and subsea applications and one aimed at low lower pluggable modules primarily for metro applications.

While the low power pluggable optics market is experiencing an influx of new suppliers, high complexity and cost in high performance optics have resulted in a small group of viable suppliers. Significantly, the high performance market has already hit the capacity limitations defined by Claude Shannon's equations. Since the introduction of 64–67 Gbaud DSPs at 600Gbps, spectral efficiency has plateaued at 8 bps/Hz and a maximum C-band capacity of 38.4Tbps, even as the latest 90+ Gbaud DSP have increased the headline channel capacity to 800Gbps.

There are three key market implications of high performance optics at the Shannon limit:

- With aggregate C-band capacity already at its maximum, new coherent generations must boost maximum per-channel data rates over longer and longer distances.
- Each new coherent generation comes at high costs and complexity for suppliers but delivers smaller gains over previous generations.
- As capacity/distance limits are reached, cost/power/efficiency metrics will become more important high performance differentiators over time.

These realities and limitations are on full display as the industry looks to the next generation of coherent DSPs beyond today's 7nm. Each supplier must individually decide whether 5nm is justified or whether it is better to skip a generation and wait for 3nm.