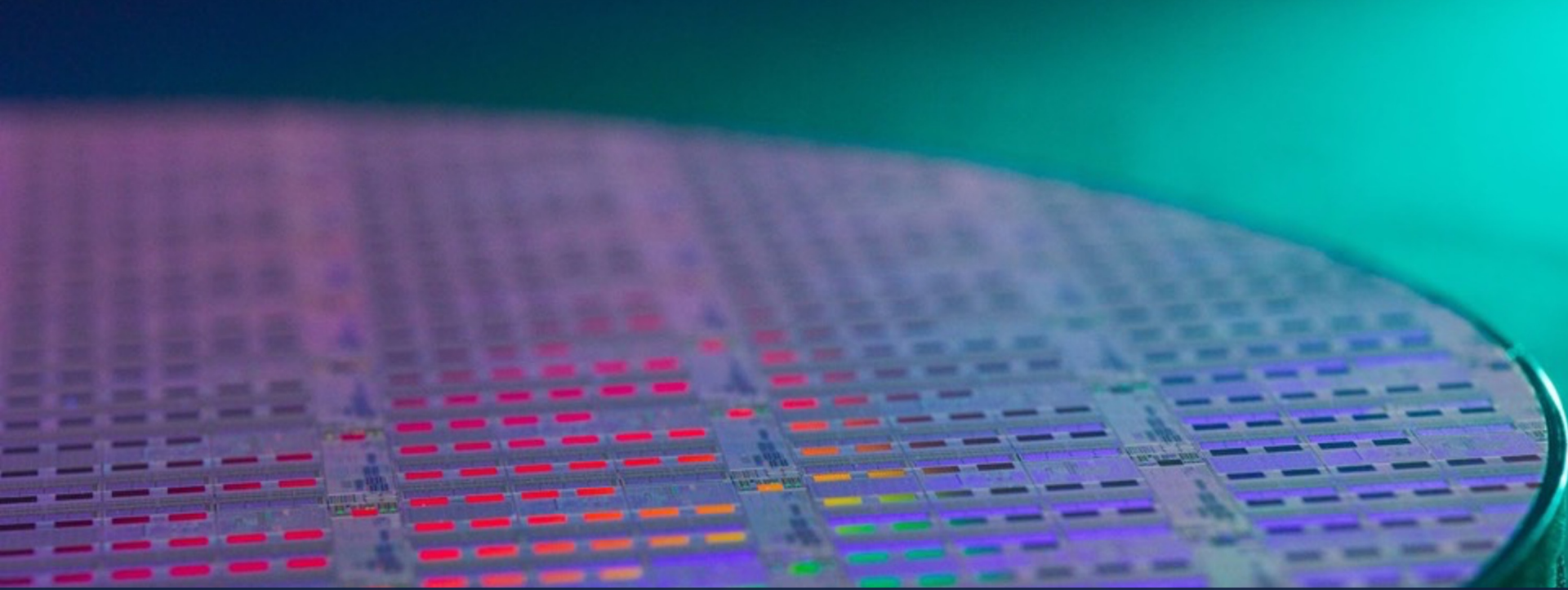


Optical I/O Technology to Meet Future Demands of HPC and AI

AyarLabs

Mark Wade, PhD | President, CTO, Co-Founder | April 25, 2023

Problem Statement for the Session

…However, disaggregation requires high levels of intra-resourcecommunication, including stringent requirements for ultra-low latency and ultra-high transmission bandwidth.

This state of the technology session poses and will explore thefollowing questions. When, where, and to what extent does disaggregation make sense for HPC systems? Will CXL, a cache-coherent interconnect for data centers, be deployed widely in HPC? Will large-scale supercomputers be disaggregated beyond rack-scale? Should we disaggregate main memory? What are the implications?

What is the state of optical I/O?

I've Heard This Story Before – What's Different?

- High value, paradigm shifting commercial application
  - Transformer based Large Language Models (GPT3, GPT4, LaMDA, LLaMA)
  - Arms race to build systems that can train larger (parameters, sequence length) models economically
- Chiplet based System-in-Package designs and advanced packaging
- New optical devices and architectures supporting multi-Tbps chips
- 300mm CMOS foundries scaling HVM (GlobalFoundries, TSMC, Intel Foundry)
- Thousands of units are already being shipped to development partners

Ayar Labs at a Glance

The Beginning

- Founded in 2015
- MIT & Berkeley research on electronics/photonics from 2010
- Built the first ever microprocessor chip with optical I/O
- Early DARPA bootstrapping

DARPA

Massachusetts Institute of Technology

Berkeley UNIVERSITY OF CALIFORNIA

Today

- Locations: Santa Clara and Emeryville CA, Boston MA
- Approximately 100 employees (85% Masters & PhD)
- 126+ patent applications filed and in process. 26 granted
- $35M+ in aggregate (DOD/DARPA, DOE, NSF) funds
- $195M of Venture Capital raised ($130M Series C Q1'22)

GlobalFoundries

FOUNDERS FUND

playground

BOARDMAN BAY CAPITAL MANAGEMENT

Si SILICON CATALYST

intel capital

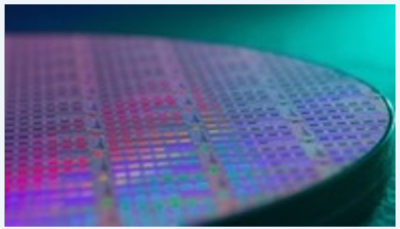
NVIDIA

APPLIED MATERIALS make possible

Hewlett Packard Enterprise

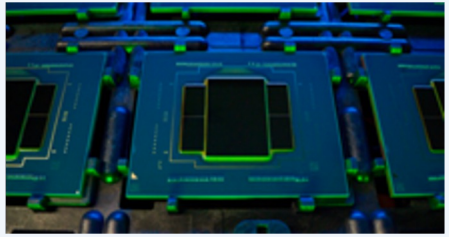
BLUE SKY CAPITAL

Technical Milestones



Photonics integrated in 300mm SOI CMOS (Ayar+GF)

Optical chiplets integrated in advanced packaging (Ayar+Intel)



Challenges to Scaling AI & HPC

Large language models (e.g. ChatGPT, Bard) are reshaping internet search - $160B/yr revenue (Google)

Strawman estimates ~$100B CapEx required to support full Google capacity[1]

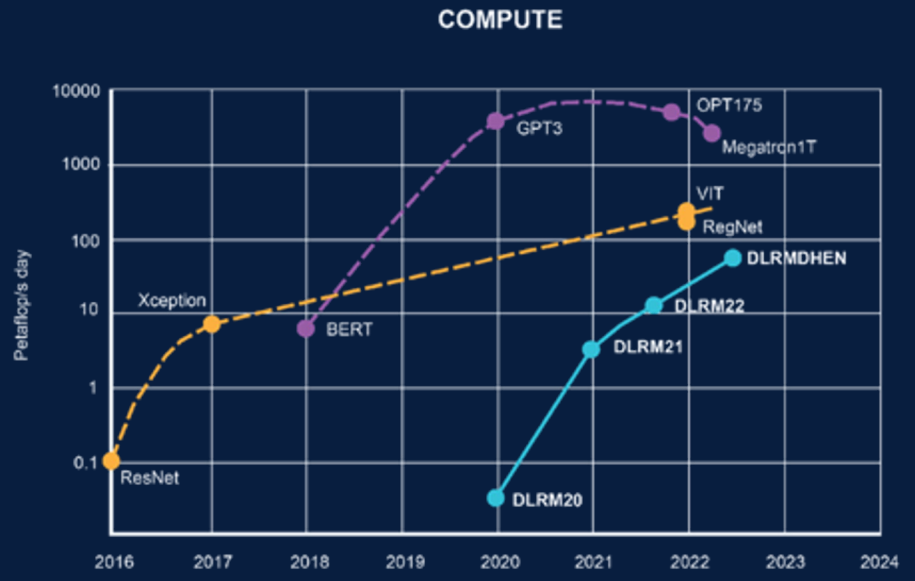
Training and inference of large models are becoming increasingly bandwidth bound (40-75% run time spent in comms)[2]

Similar distributed computing system challenges between AI & HPC – Exascale efficiencies result in 500MW projected for Zettascale[3]

Advanced packaging and heterogeneous integration enables optical I/O chiplets – significantly changing the traditional bandwidth-distance constraints

[1] https://www.semianalysis.com/p/the-inference-cost-of-search-disruption
[2] Pati, et al, Computation vs. Communication Scaling for Future Transformers on Future Hardware, https://arxiv.org/ftp/arxiv/papers/2302/2302.02825.pdf
[3] Lisa Su, ISSCC Plenary, 2023

Machine Learning Trends

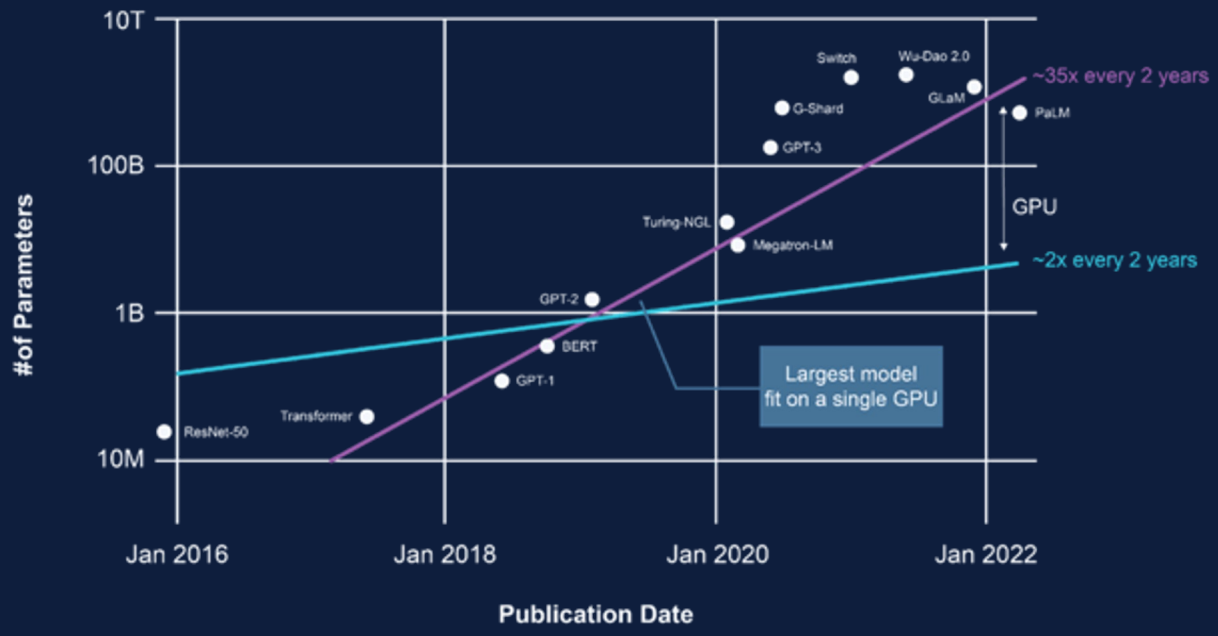


[Source: Meta OCP Global Summit 2022]

- 10,000x growth in model size and compute requirements in ~5 years
- ~$10M energy bill to train one model
- Insatiable model growth (parameter size, sequence lengths) create tremendous hardware strain

Model Growth Outpacing Hardware

Growing gap between memory demand and supply



[Source: NVIDIA COBO Workshop Nov 2022]

- Largest model that can fit on one GPU is ~1-10B parameters
- Getting to >>10B parameter size models requires parallelizing the models across many sockets (i.e. scale-out)
- Scale-out architectures create tremendous pressure on the communications fabric

OpenAI

ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response.

(Released in December 2022)

"OpenAI must be on the cutting edge of AI capabilities and low latency, high bandwidth optical interconnect is a central piece of our compute strategy to achieve our mission of delivering artificial intelligence technology that benefits all of humanity."

- Chris Berner
Head of Compute
OpenAI

Forbes

FORBES > MONEY

Microsoft Confirms Its $10 Billion Investment Into ChatGPT, Changing How Microsoft Competes With Google, Apple And Other Tech Giants

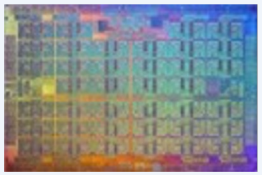
Q.ai - Powering a Personal Wealth Movement

"OpenAI must be on the cutting edge of AI capabilities and low latency, high bandwidth optical interconnect is a central piece of our compute strategy to achieve our mission of delivering artificial intelligence technology that benefits all of humanity."

- Chris Berner
Head of Compute
OpenAI

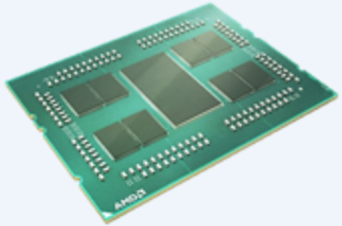
Optical I/O can Redefine the compute “Socket”

Past



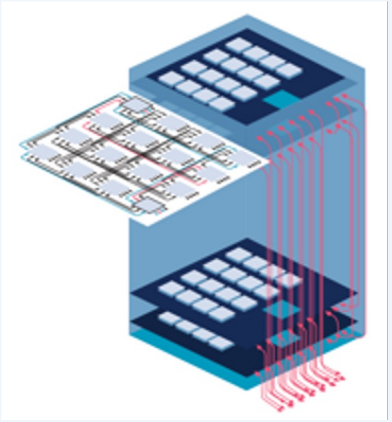
Intel® Xeon Phi
8 Billion Transistors

Present



AMD's 64-core EPYC CPU
~40 Billion Transistors

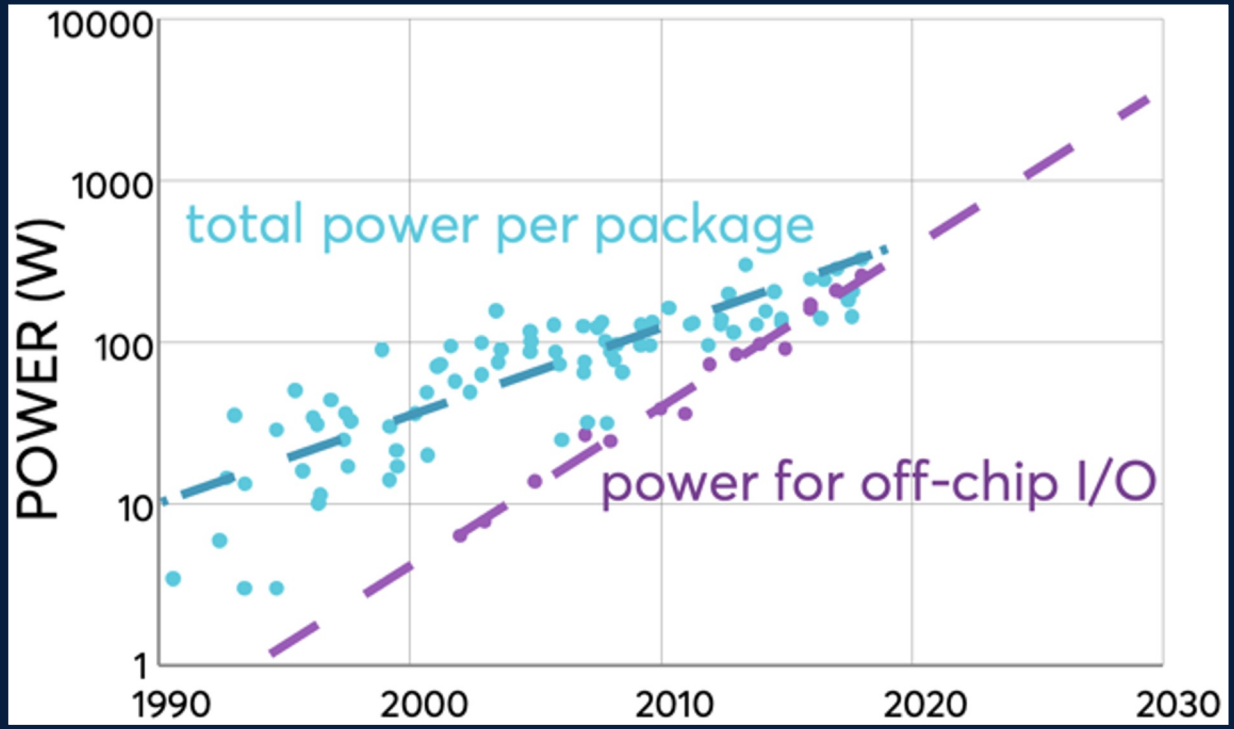
Future with Optical I/O



The Rack is the Socket
20+ Trillion Transistors

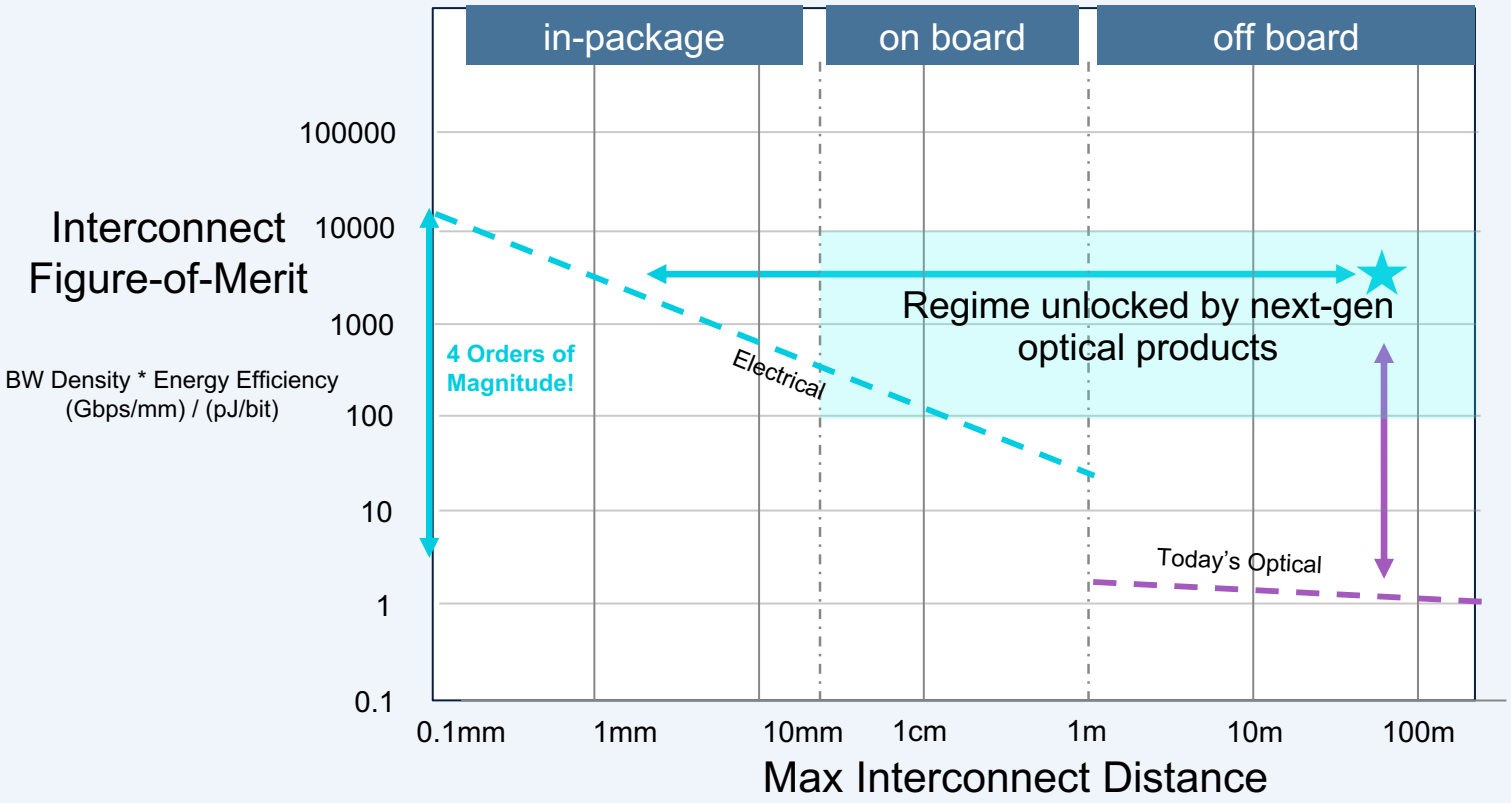
CPU’s are many compute cores and functions wrapped in a power efficient, low latency, high bandwidth interconnect. Optical I/O has these characteristics but with extended reach

The Challenge: The Bandwidth Bottleneck Power Wall



[G. Keeler, DARPA ERI Summit 2019]

The Challenge: Electrical Signaling is Not Scaling



[G. Keeler, DARPA ERI Summit 2019]

Advanced Chiplet Packaging Enables Optical I/O

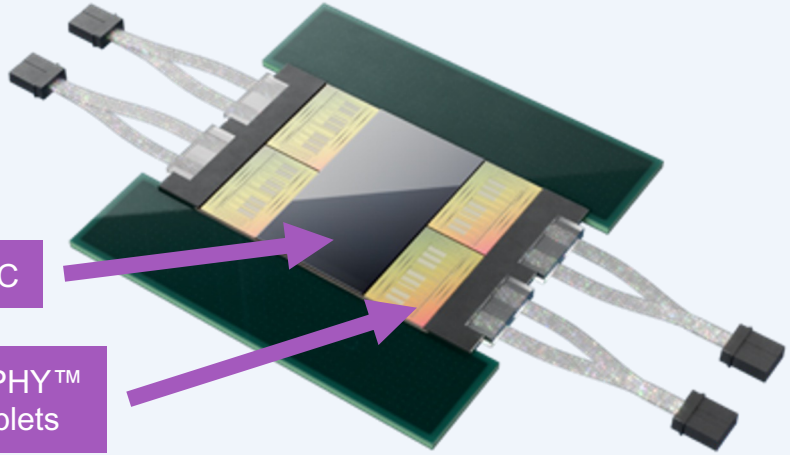
Intel Ponte Vecchio GPU
(Source: Intel)

AMD MI300 (CPU+GPU)
(Source: AMD)

Nvidia Hopper GPU
(Source: Nvidia)

CPU/GPU/FPGA/SOC

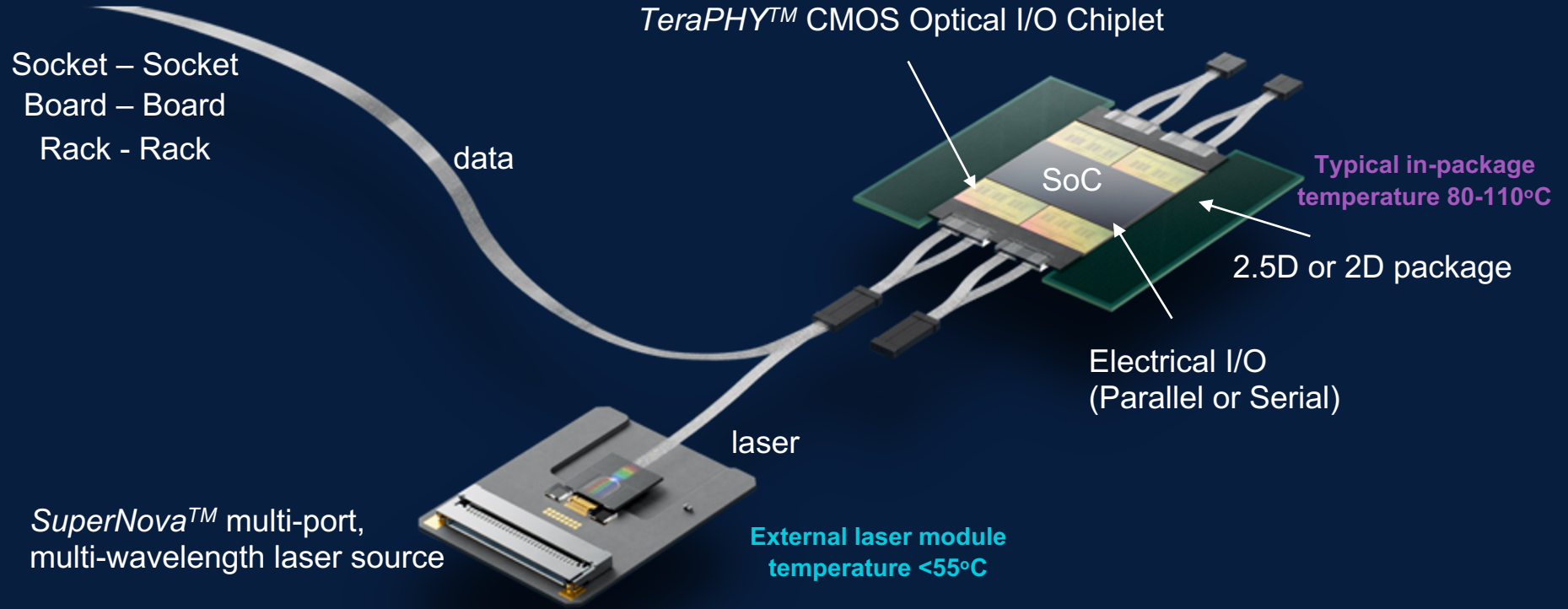
Ayar Labs TeraPHY™
Optical I/O chiplets



Chiplet Integration Platforms

- Leverage existing CMOS foundry processes
- Compatible with downstream OSAT capabilities
- Intercept emerging advanced chiplet packaging technologies

Ayar Labs Optical I/O Solution



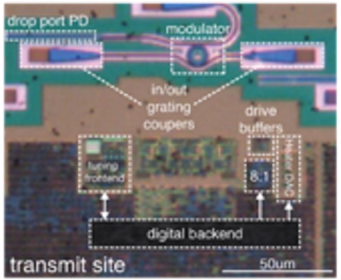
The Ayar Labs Optical I/O solution breaks the bandwidth-distance bottleneck

Technology Basics

Microring Resonators

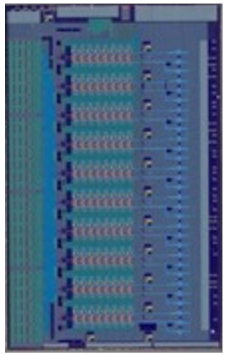
- 1,000x smaller than optical devices
- High-speed capability
- Compatible with 300mm CMOS

Electronic/Photonic Integration



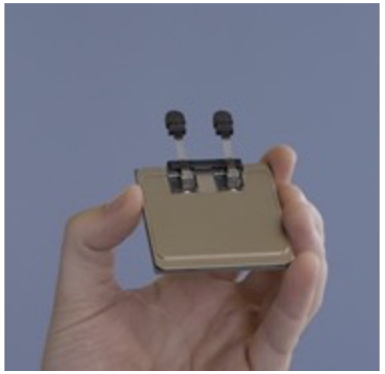
- Dense CMOS integration

Optical chiplets



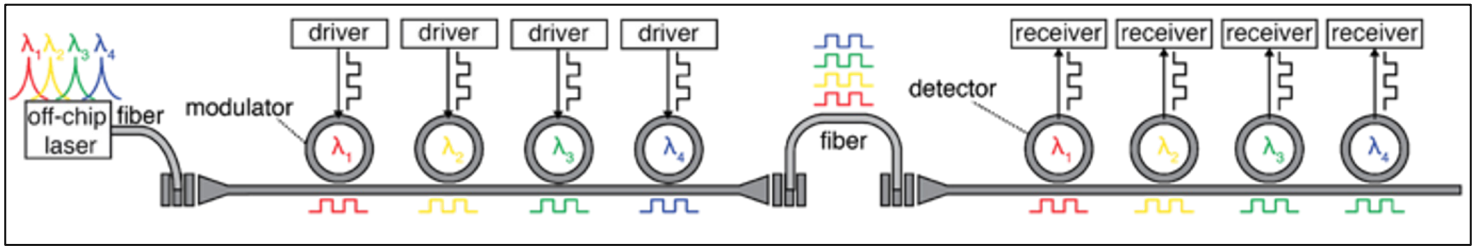
- TeraPHY™ chiplet for in-package optical I/O

SoC In-Package Integration



- Integration with state-of-the-art SoCs
- Direct from the package optical I/O

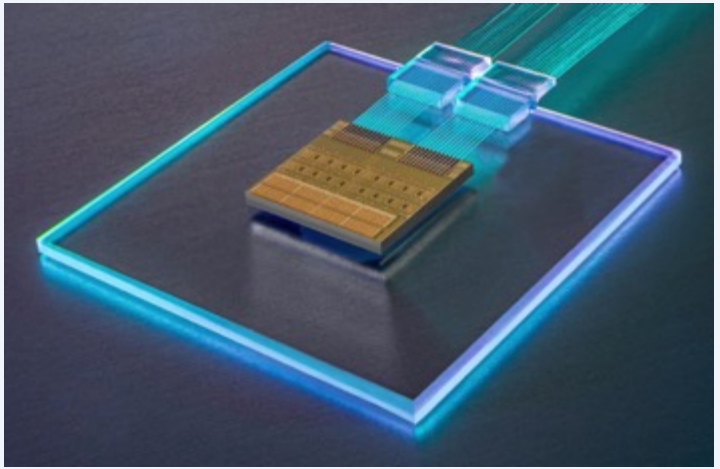
Microring WDM Bandwidth Scaling (Tx+Rx)



Chiplet bandwidth = 2 * (# of ports/chiplet) x (# of wavelengths/port) x (data rate/wavelength)

Chiplet Bandwidth	# of ports / chiplet	# of wavelengths/port	Data rate/wavelength
4.096 Tbps	8	8	32 Gbps
8.192 Tbps	16 (8)	8	32 Gbps (64 Gbps)
16.384 Tbps	16	8	64 Gbps
32.768 Tbps	16	16	64 Gbps

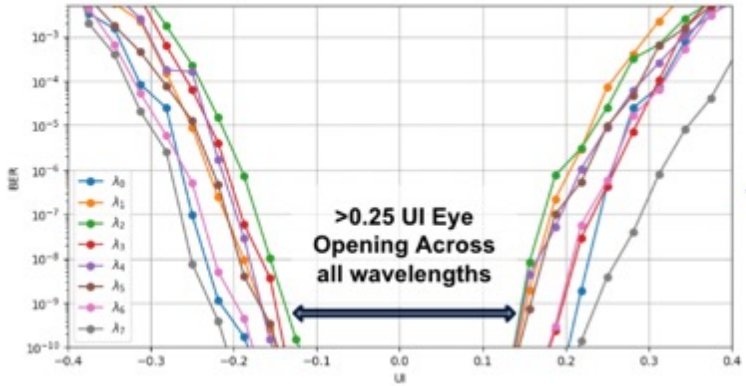
Publicly Demonstrating Products



Presented at Optical Fiber Conference (OFC) 2023

Live Demonstration of Industry’s first 4-Tbps Optical Solution

Industry First 4 Tbps Optical I/O Demonstrations

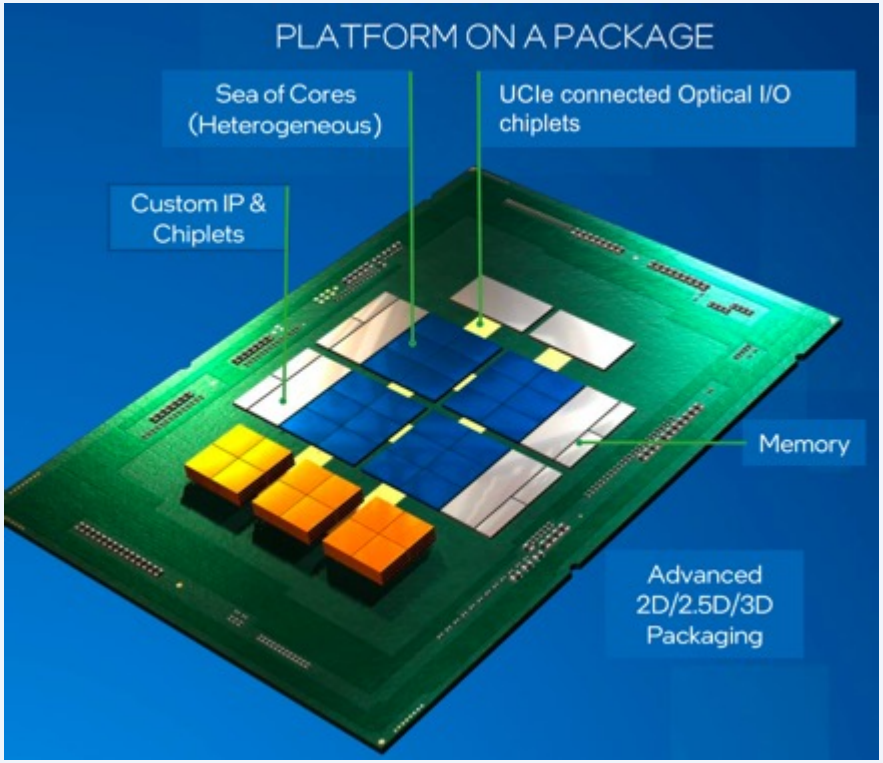


AyarLabs

System Summary
Link Summary
Margin View
Laser Control

	Link Name	TX Macro	TX Lock	RX Macro	RX Lock	Bits	BER
0	8x32Gbps_L2R_0	0	✅	0	✅	1.7027e+15	1.8794e-14
1	8x32Gbps_R2L_0	0	✅	0	✅	1.8589e+15	0.0000e+00
2	8x32Gbps_L2R_1	1	✅	1	✅	1.7944e+15	0.0000e+00
3	8x32Gbps_R2L_1	1	✅	1	✅	1.6994e+15	0.0000e+00
4	8x32Gbps_L2R_2	2	✅	2	✅	1.8589e+15	0.0000e+00
5	8x32Gbps_R2L_2	2	✅	2	✅	1.8453e+15	0.0000e+00
6	8x32Gbps_L2R_3	3	✅	3	✅	1.6177e+15	0.0000e+00
7	8x32Gbps_R2L_3	3	✅	3	✅	1.8015e+15	1.6655e-15
8	8x32Gbps_L2R_4	4	✅	4	✅	1.6798e+15	0.0000e+00
9	8x32Gbps_R2L_4	4	✅	4	✅	1.8266e+15	2.2994e-14
10	8x32Gbps_L2R_5	5	✅	5	✅	1.8585e+15	0.0000e+00
11	8x32Gbps_R2L_5	5	✅	5	✅	1.7453e+15	0.0000e+00
12	8x32Gbps_L2R_6	6	✅	6	✅	1.8393e+15	0.0000e+00
13	8x32Gbps_R2L_6	6	✅	6	✅	1.6315e+15	7.3552e-15
14	8x32Gbps_L2R_7	7	✅	7	✅	1.7381e+15	0.0000e+00
15	8x32Gbps_R2L_7	7	✅	7	✅	1.7777e+15	3.3751e-15

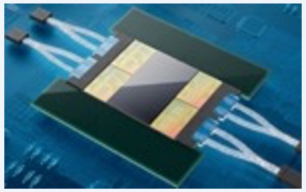
Future Systems In Package with Optical I/O



(Source: Modified from UCIe spec, Intel, AMD)

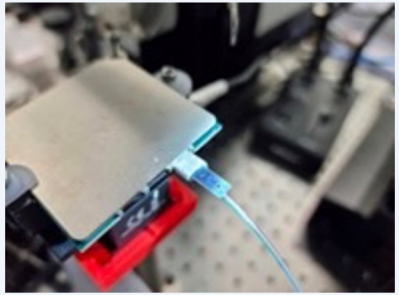
Gen	Electrical I/F (Advanced Package)				Optical I/F (CW-WDM)			Optical Chiplet BW (Tx+Rx)	Off-package IO BW (4-8 chiplets per package)
	I/F	Modules	Tx / Rx IOs	Data Rate [Gbps/IO]	Ports	λs / Port	Data Rate [Gbps/λ]		
1	AIB	24	20 / 20	2	8	8	16	2 Tbps	8-16 Tbps
2	AIB	16	80 / 80	2	8	8	32	4 Tbps	16-32 Tbps
3	UCIe	16	32 / 32	8	8	16	32	8 Tbps	32-65 Tbps
4	UCIe	16	64 / 64	8	16	16	32	16 Tbps	65-131 Tbps
5	UCIe	16	64 / 64	16	16	16	64	32 Tbps	131-262 Tbps

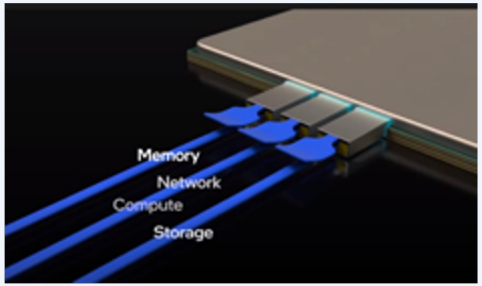
- Gen 1 and Gen 2 already built and hardware validated
- 16-32 Tbps off-socket optical I/O bandwidth possible today
- Clear multi-generation roadmap leveraging advanced packaging and industry standards
- >250 Tbps off-socket optical I/O bandwidth possible in 10-15 year time frame



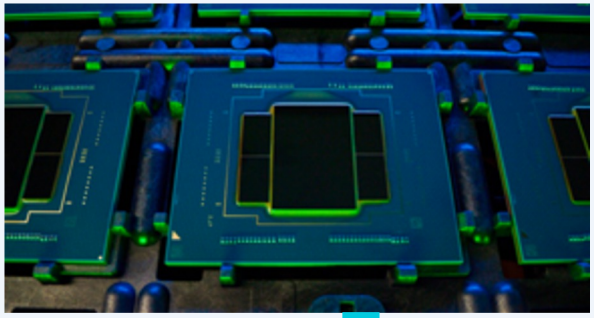
Packaging, Fiber Attach and System Integration

Package Level Pluggable Optical Connectors





(Intel Innovation Day 2022)

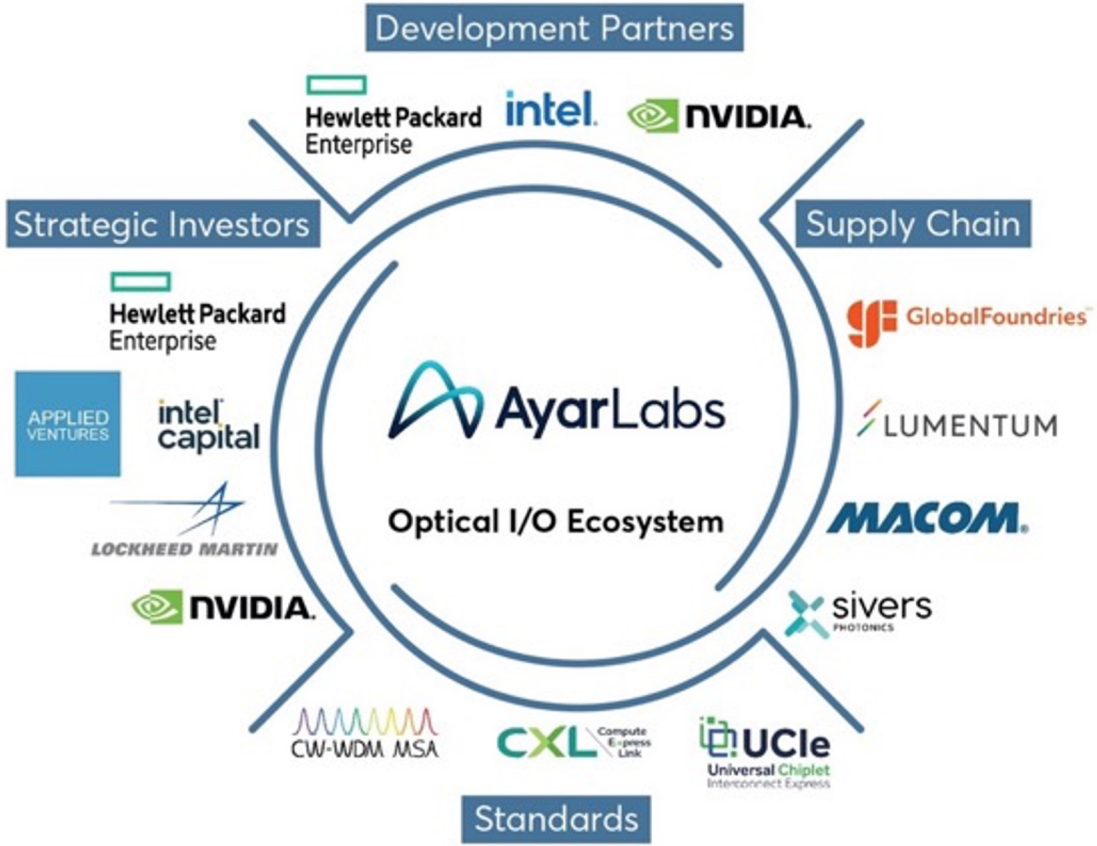


Multi-chip packages with optical chiplets assembled into standard form factors

(Intel + Ayar OFC 2021)

(Intel + Ayar OFC 2023)

Go-to-Market Ecosystem



Partnering across the HVM ecosystem

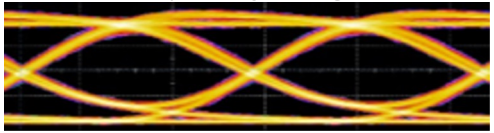
Path to Production: Recent Progress

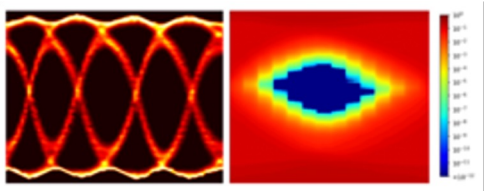
Completed Product Validation

4 Tbps (Tx+Rx)

	Link Name	TX Macro	TX Lock	RX Macro	RX Lock	Bits	BER
0	olink_D2-520_b_0	0	✓	1	✓	1.1377e+16	2.0213e-15
1	olink_D2-520_b_1	1	✓	0	✓	1.1332e+16	6.1773e-15
2	olink_D2-520_b_2	2	✓	2	✓	1.1301e+16	8.8492e-15
3	olink_D2-520_b_3	3	✓	6	✓	1.1277e+16	3.8129e-15
4	olink_D2-520_b_4	4	✓	4	✓	1.1238e+16	2.6714e-16
5	olink_D2-520_b_5	5	✓	5	✓	1.1175e+16	6.2640e-16
6	olink_D2-520_b_6	6	✓	3	✓	1.1120e+16	1.7985e-16
7	olink_D2-520_b_7	7	✓	7	✓	1.1052e+16	6.3336e-16

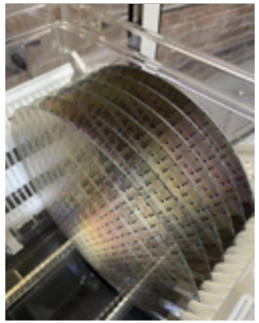
Sample TX eye diagram @ 32Gbps

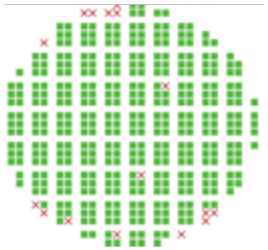




Sample RX eye diagram & BER sweep

Established Manufacturing Line and Currently Shipping

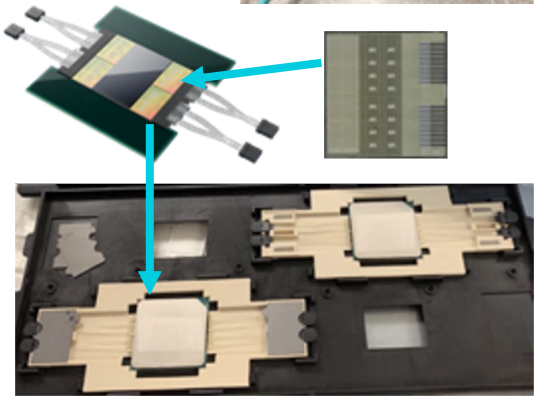




Established KGD methodology

Customer Platform Integration

2022 fully-assembled hardware, performance validated



Platform bring-up happening now

Status Check of Optical I/O – is it ready?

Does it work?

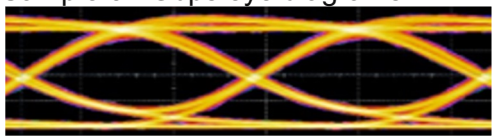
Is it manufacturable?

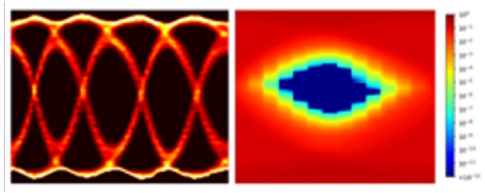
Does the cost structure scale?

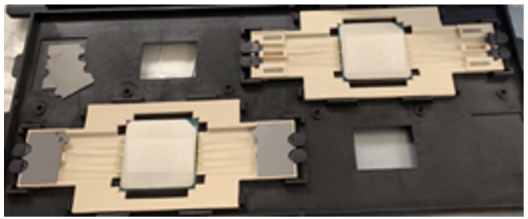
Status Check of Optical I/O – is it ready?

Does it work?

Sample 32 Gbps eye diagrams



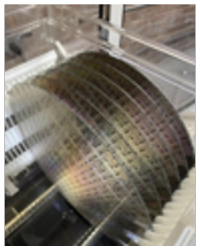




Integration into early customer adoption
16 Tbps off-socket BW

- 2.2x higher than Nvidia H100
- Equivalent to 256 lanes of PCIe Gen5
- <1e-12 native BER (no heavy FEC needed)

Is it manufacturable?



Does the cost structure scale?

Already shipping thousands of engineering sample units

Cost structure drivers:

1) 300mm HVM CMOS economies of scale
2) Laser die and modules designed and assembled with HVM partners
3) Increased integration of optical functionality
4) A single laser is shared across many optical channels

Thank You!