Optics in Future AI Systems: Interconnects, Switching and Processing

Open Compute Project

Panel Discussion at OFC

Tuesday, 07 March, 10:45 – 11:45

Theater II

OPEN Compute Project®

Community-driven hyperscale innovation for all.

Optics for Future AI Systems: Interconnects, Switching and Processing

Vlad Kozlov	**Andreas** Bechtolsheim	**Loi** Nguyen	**Uday** Poosarla	**Hamid** Arabzadeh
				
CEO, LightCounting	**OCP Board, Co-Founder and Chairman, Arista**	**EVP/GM, Optical and Copper Connectivity Group Marvell**	**Technical Sourcing Manager, Optical Interconnects, Meta**	**Chairman and CEO, Ranovus**
	Extra Content	Extra Content	Extra Content	Extra Content

Vlad Kozlov

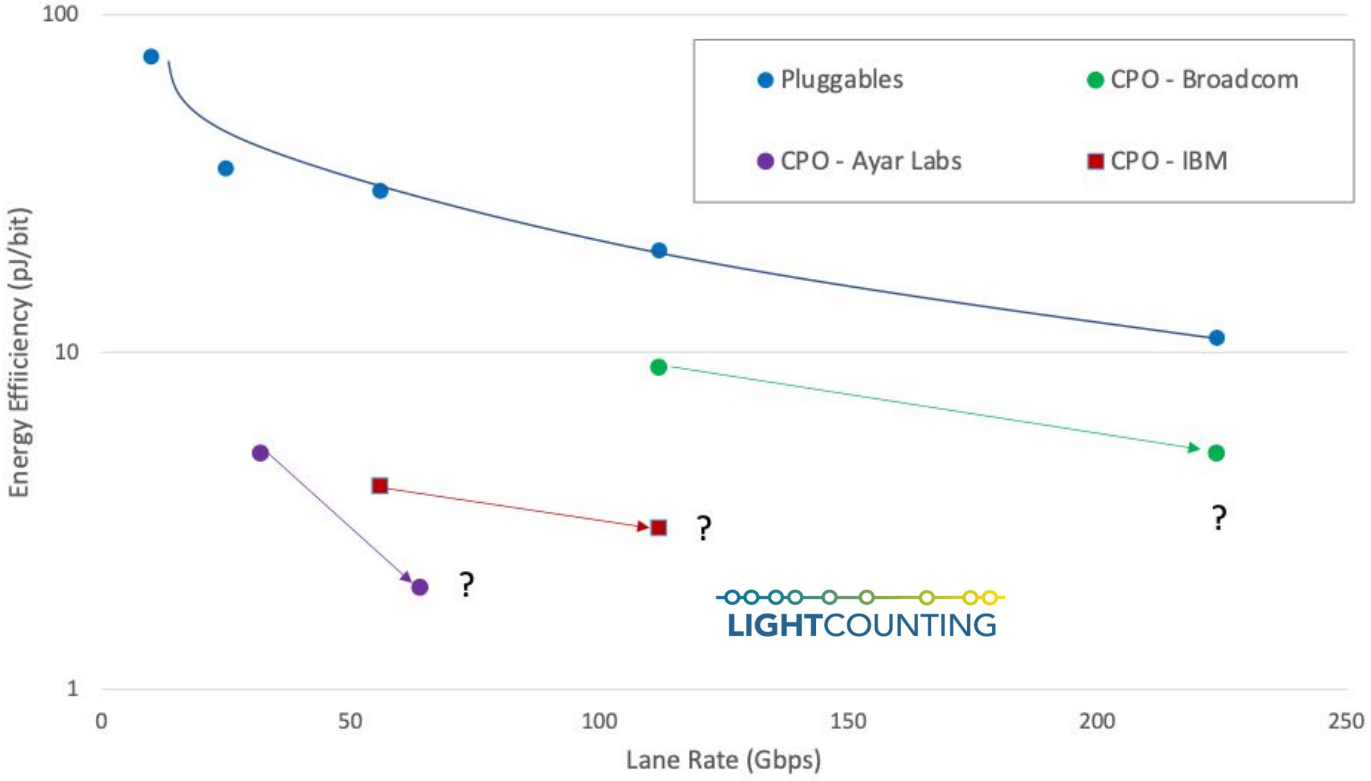
CEO
LightCounting

Open Possibilities.

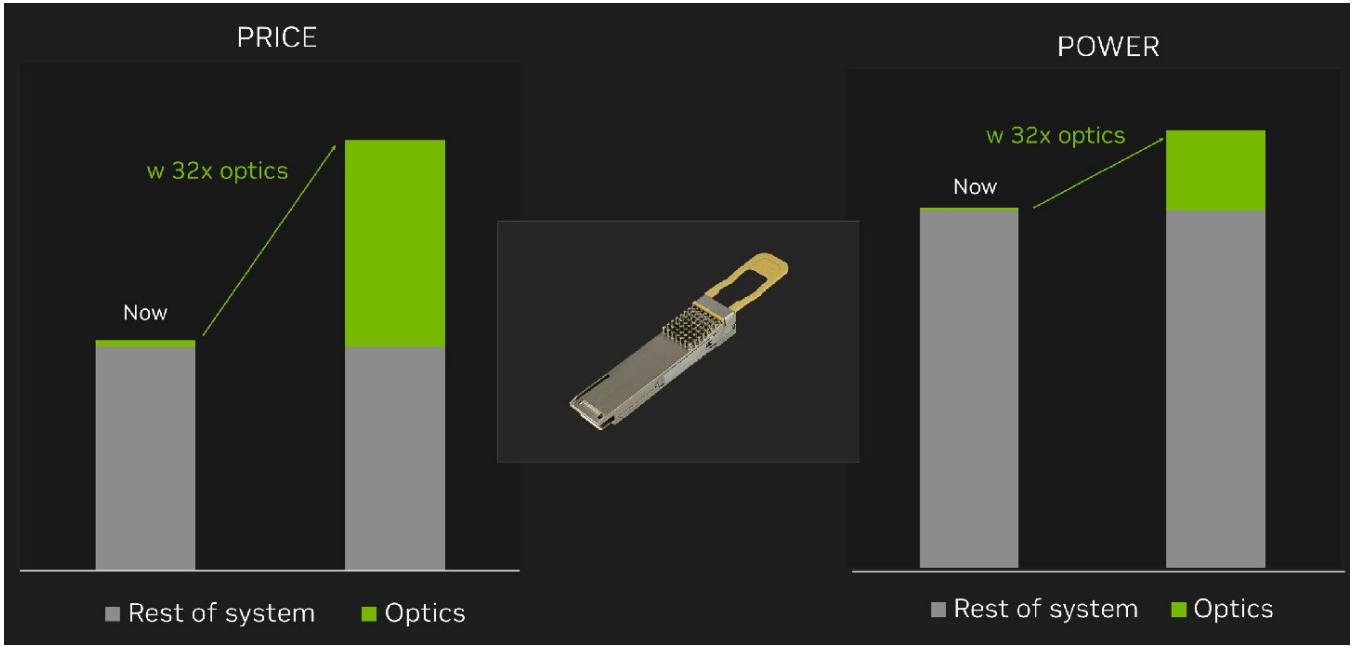
YEARS

OPEN Compute Project®

Progress in reducing power consumption

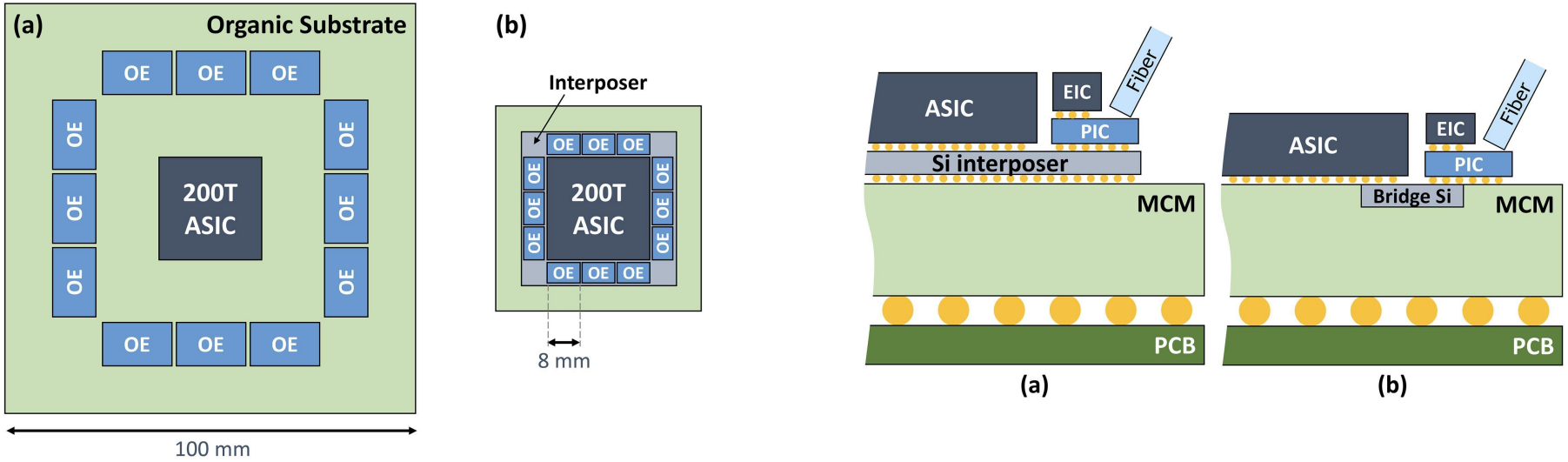


NVIDIA needs 32x more bandwidth now



Source:
NVidia

Beyond CPO: 1pJ/bit and >1Tbps/mm



Source: Nvidia, Beyond CPO: A Motivation and Approach for Bringing Optics Onto the Silicon Interposer, **published Feb 15, 2023**

Thank You

Home

OPEN Compute Project®

Community-driven hyperscale innovation for all.

Uday Poosarla

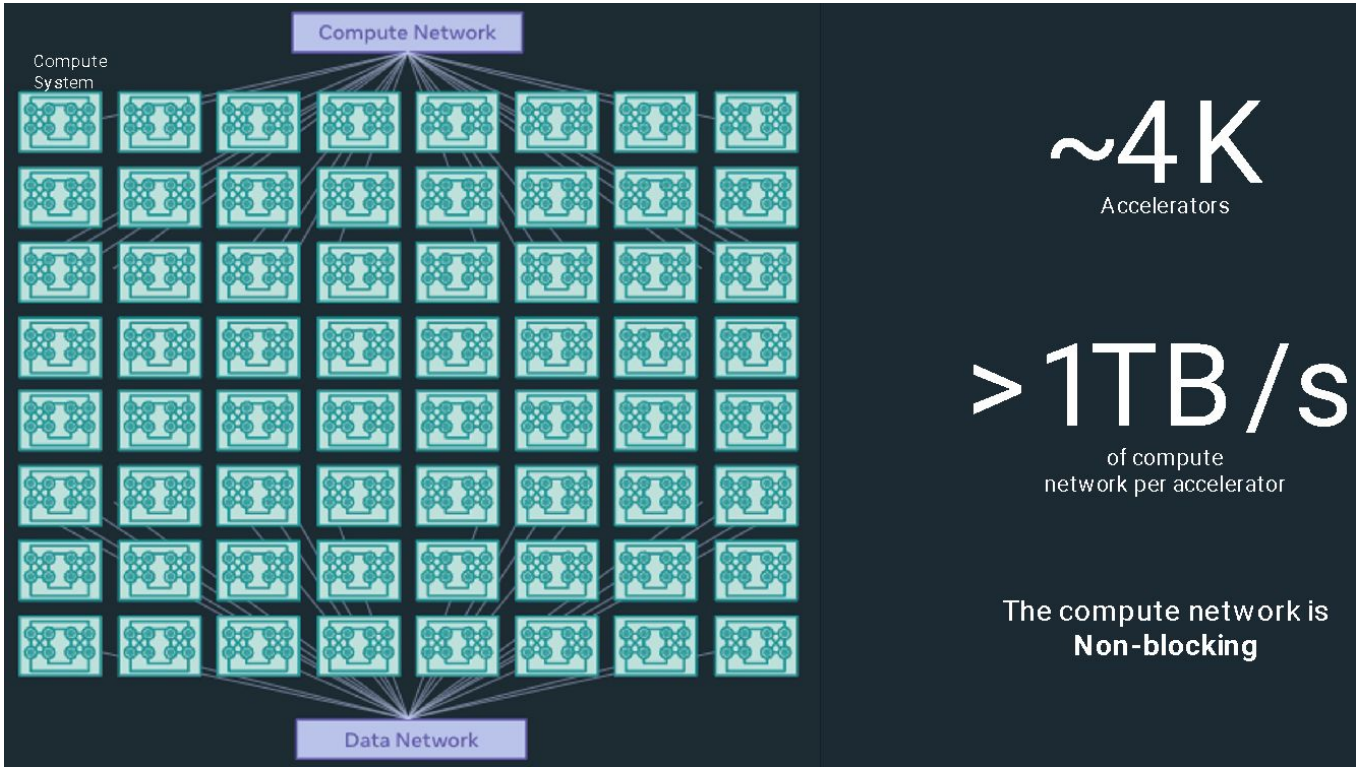
Technical Sourcing Manager, Optical Interconnects, Meta

Open Possibilities.

10 YEARS

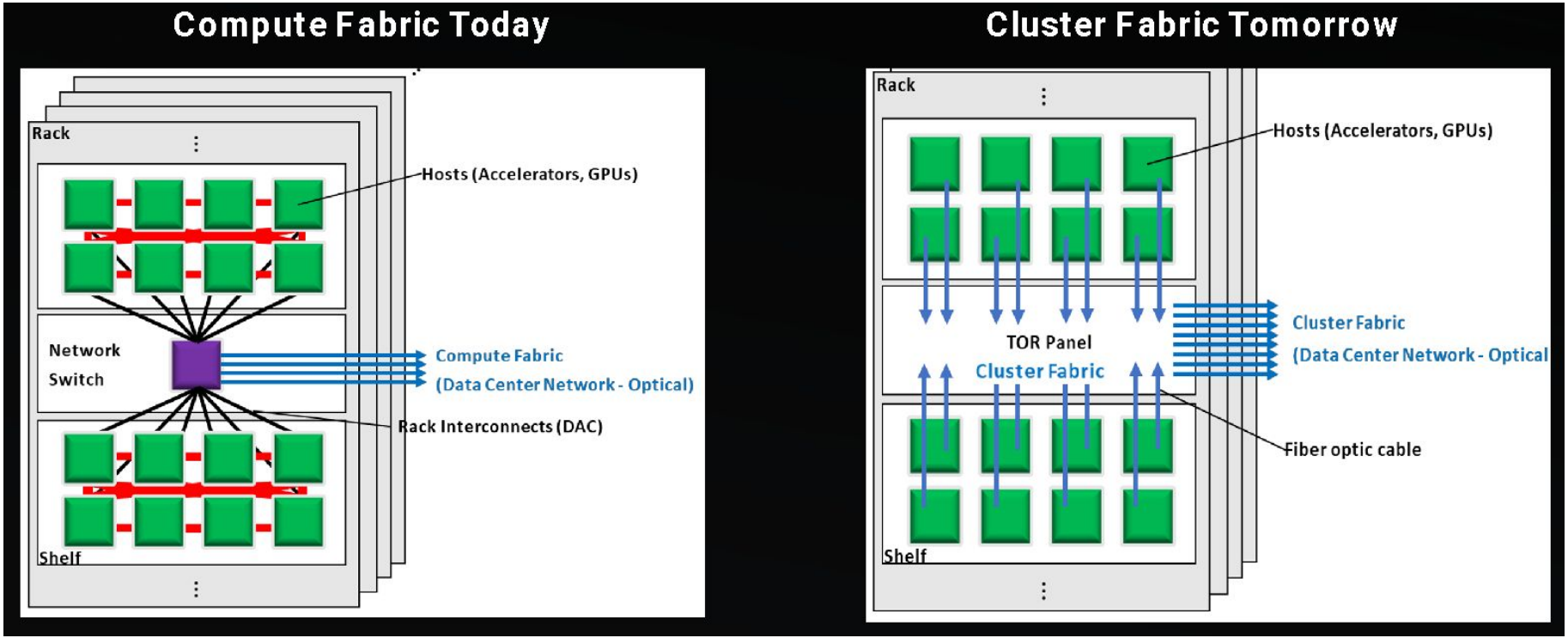
OPEN Compute Project®

AI Training Cluster of the Future (2025+)



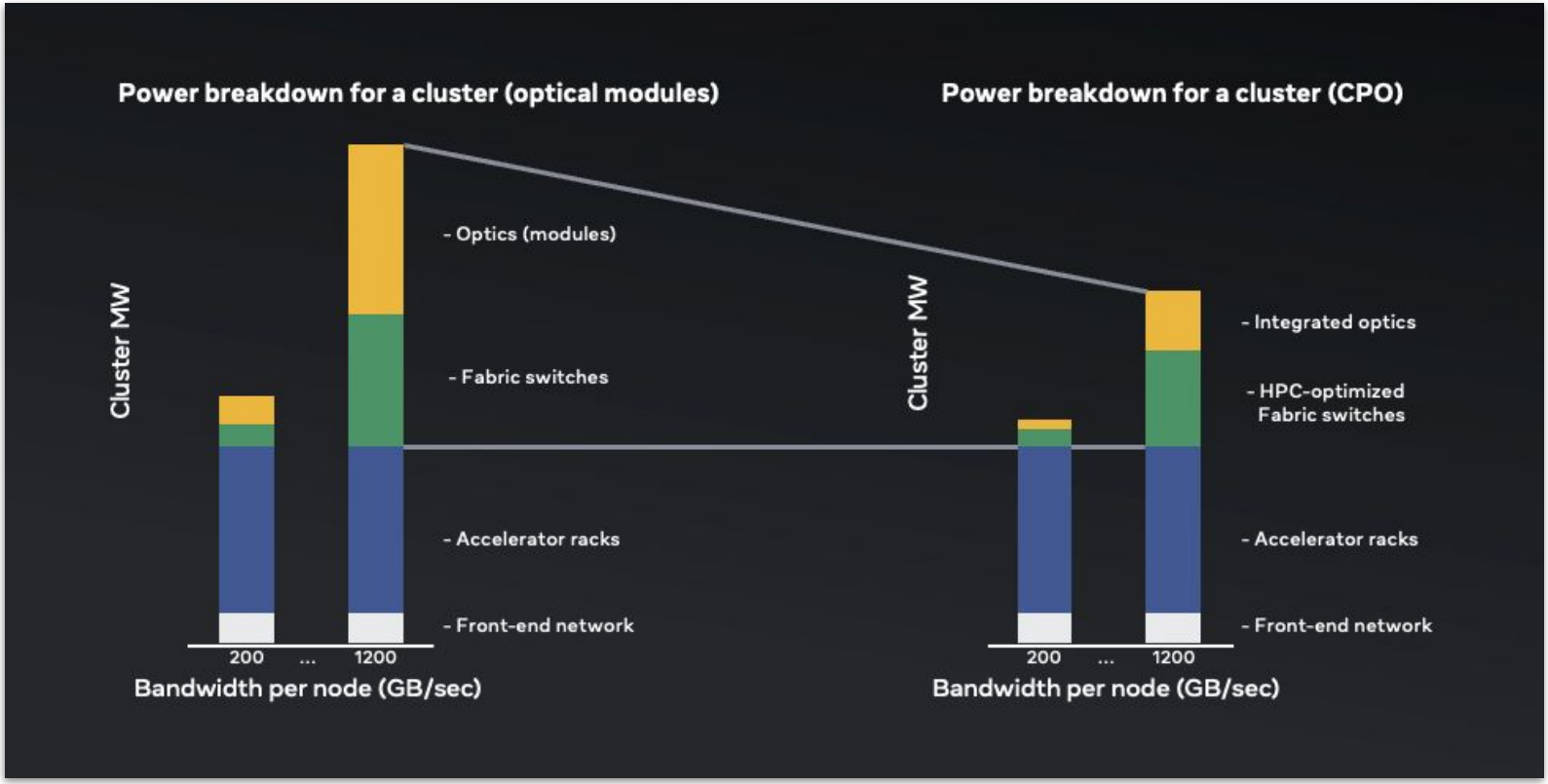
Source: Meta

Interconnect Scaling Challenges

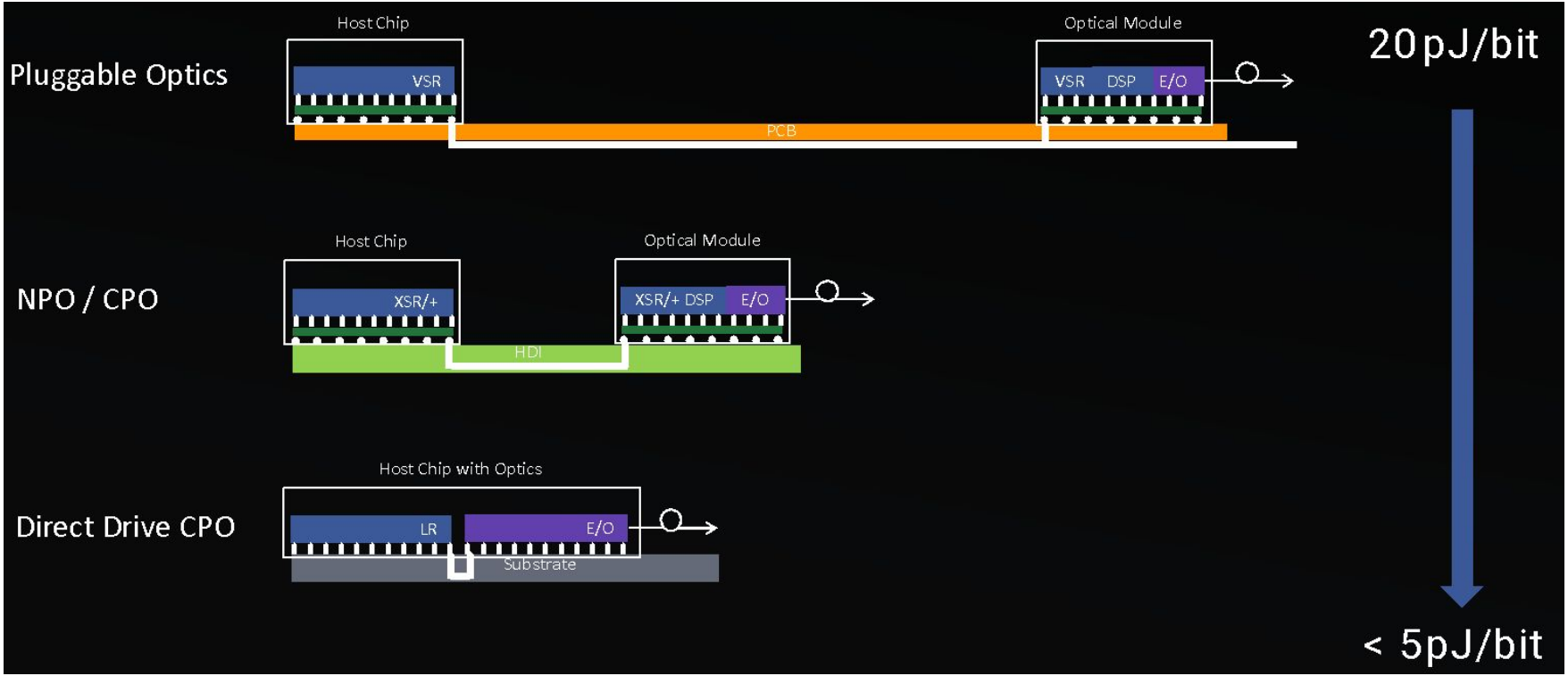


Source: Meta

Optical I/O


Power breakdown for a cluster (optical modules)
Cluster MW
- Optics (modules)
- Fabric switches
- Accelerator racks
- Front-end network
200
...
1200
Bandwidth per node (GB/sec)
Power breakdown for a cluster (CPO)
Cluster MW
- Integrated optics
- HPC-optimized Fabric switches
- Accelerator racks
- Front-end network
200
...
1200
Bandwidth per node (GB/sec)

Integration has Power Advantage



Source: Meta

Thank You

Home

OPEN Compute Project®

Community-driven hyperscale innovation for all.

Andreas Bechtolsheim

OCP Board, Co-Founder and Chairman
Arista

Open Possibilities.

10 YEARS
OPEN Compute Project®

Linear Drive Optics Modules

1. **Linear Drive means no DSP or CDR**

   Just a linear driver to provide required modulator voltage

2. **Requires a high-performance switch SERDES**

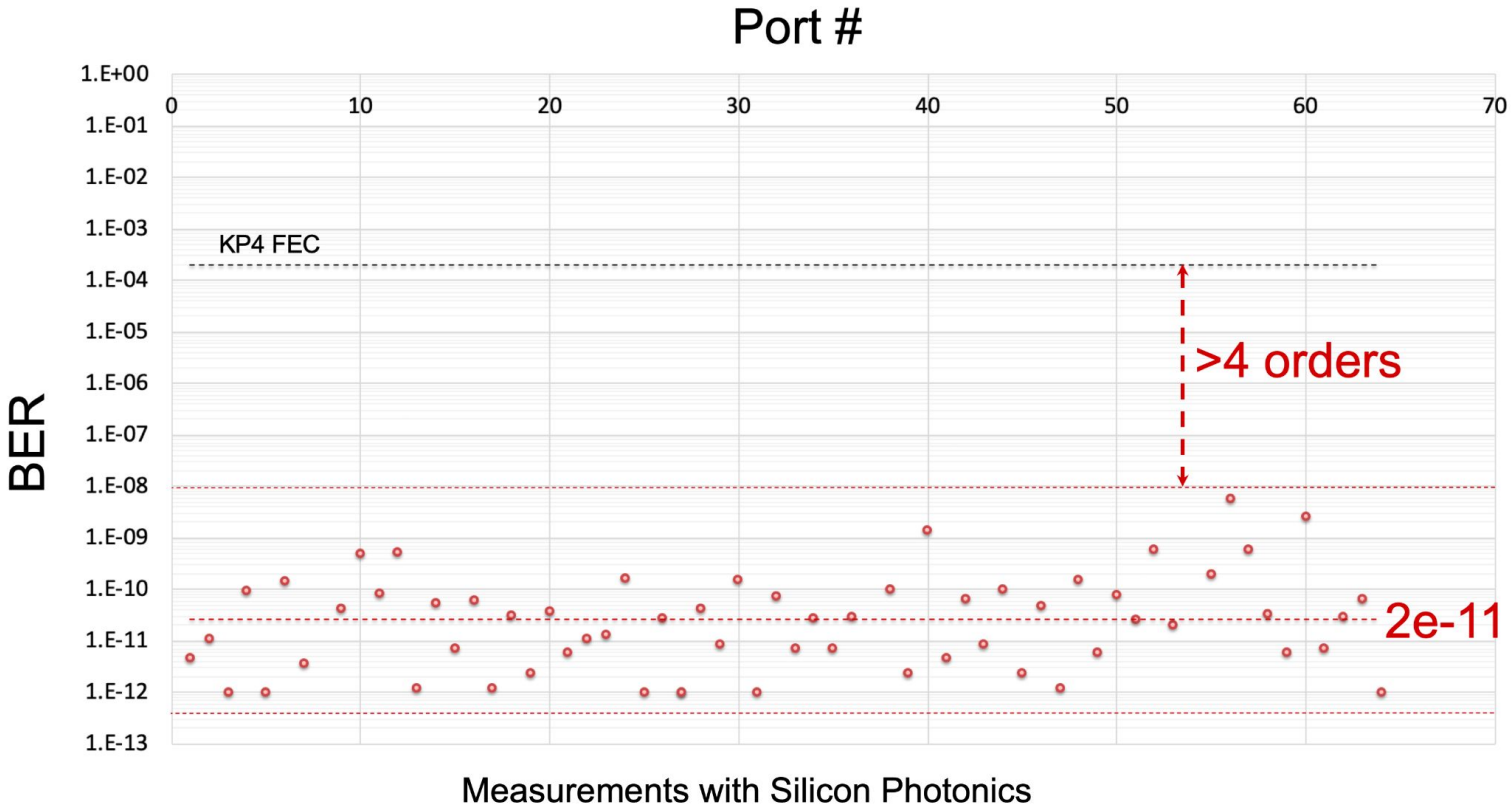
   And very careful signal integrity design

3. **Achieves power savings similar to direct drive CPO**

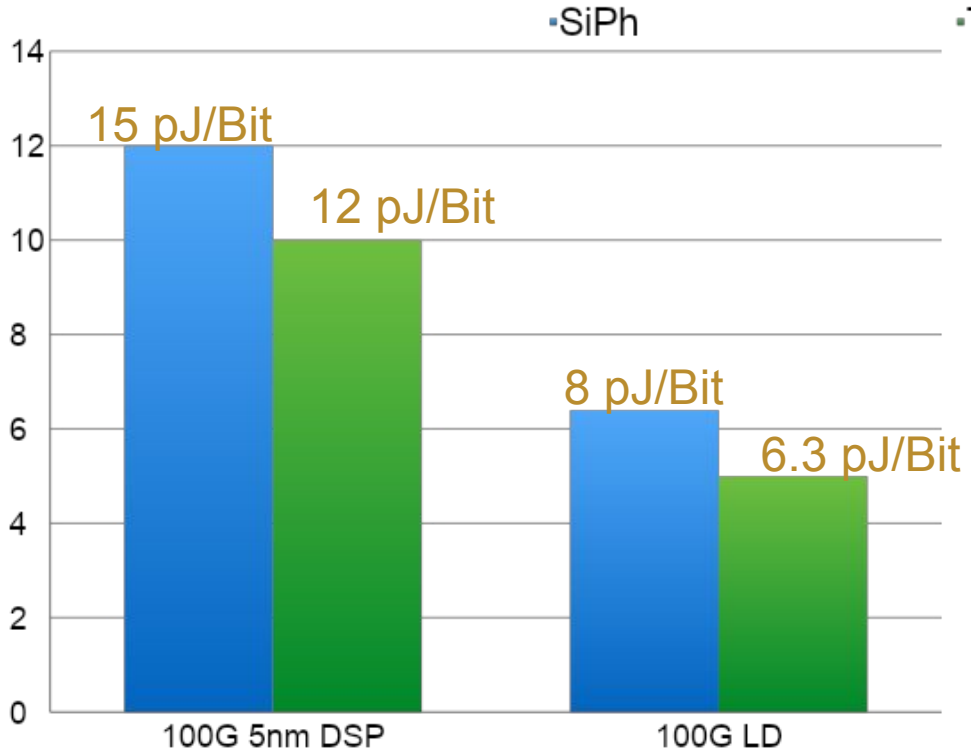
   While retaining the many advantages of pluggable optics modules

Opportunity to cut optics module power by 50%
and system power by up to 25%

Pre-FEC BER with Ethernet Traffic



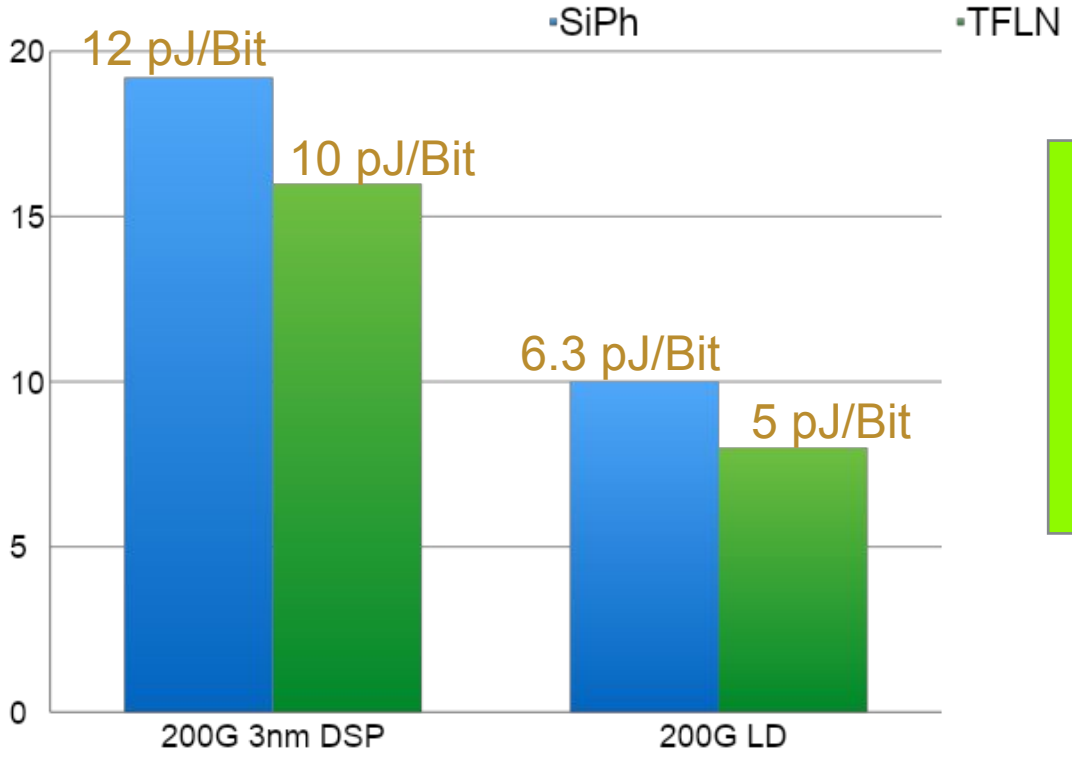
800G LD Pluggables Power Evolution



Measured Power at 3.3V
for 800G SiPh and TFLN
Pluggable Optics Modules

pJ/Bit Calculated at 100G
All Power is typical

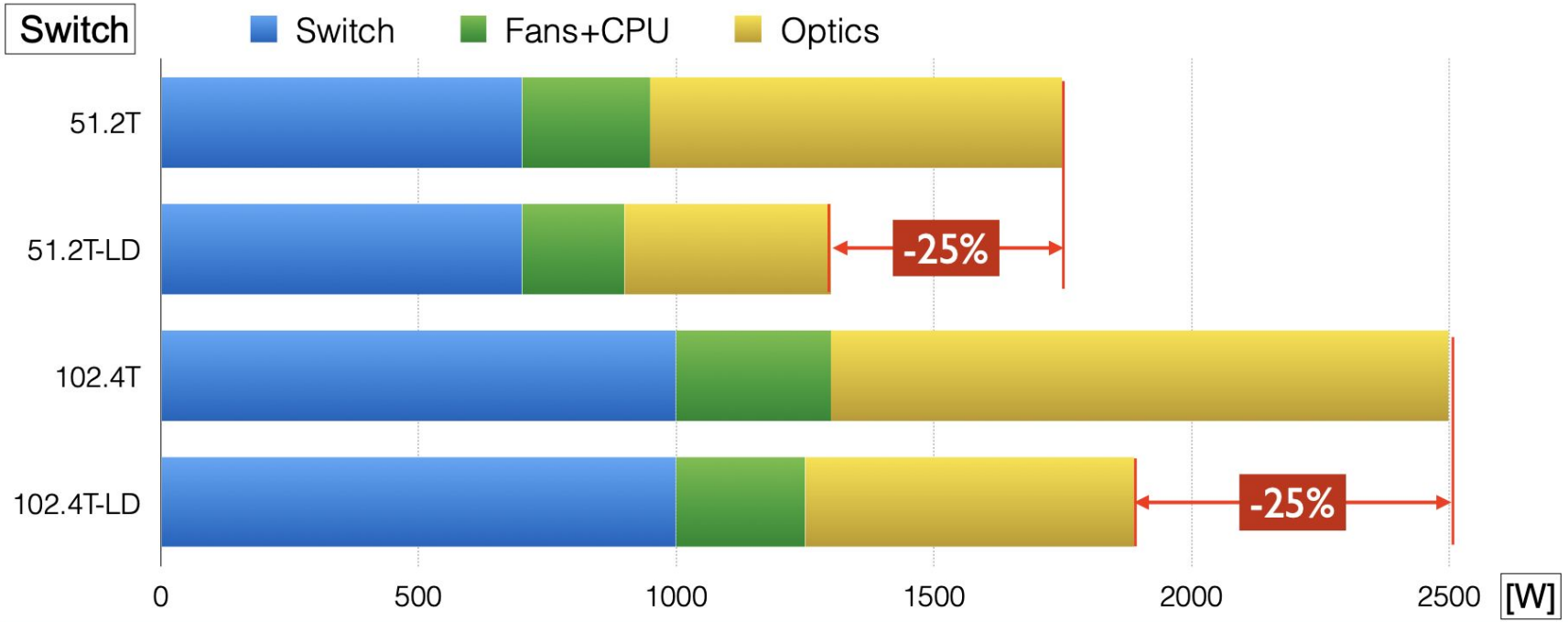
1600G LD Pluggables Power Evolution



Projected Power at 3.3V
for 1600G SiPh and TFLN
Pluggable Optics Modules

pJ/Bit Calculated at 200G
All power is typical

System Level Power Savings with LD Optics



Additional power savings possible with non SiPh Modulators

EOM Technology Comparison for 200G

Technology	Integration Capability	Bandwidth	Vpi (1)	Insertion Loss (2)	Reliability	HVM (3)
Silicon Photonics	Excellent	Sufficient	High	High	Proven	Now
III-V	Low	Higher	Low	Low	Proven	Now
TFLN	Low	Very High	Low	Low	Proven	2024 (E)
BTO	New Process	Very High	Low	Low	Proven	2025 (E)
Organic	New Process	Very High	Low	Low	To be Proven	2026 (E)

(1) Lower Vpi lowers driver power consumption
(2) Lower insertion loss lowers laser power
(3) HVM = High volume manufacturing

LPO Comparison to Direct Drive CPO at 200G λ

Form Factor	LPO	CPO	CPO Delta
DSP	N/A	N/A	Same
Linear Driver	Required	N/A	Lower Power
External Laser	N/A	Required	Higher Power
New EOM Tech	Supported	Not POR	Higher Power
Voltage Rail	3.3V	Device Voltage	Adjust for voltage
Projected Power	6.3 pJ/Bit SiPh 5 pJ/Bit TFLN	5 pJ/Bit SiPh @ Device Voltage	Very similar

Laser Efficiency at Temperature

1. Laser efficiency at high temperature is key

Laser temperature inside modules can reach 75C - 85C

2. Desirable to avoid Thermo-Electric Coolers

TECs significantly reduce net laser power efficiency

3. QD Lasers offer highest efficiency at temp without TEC

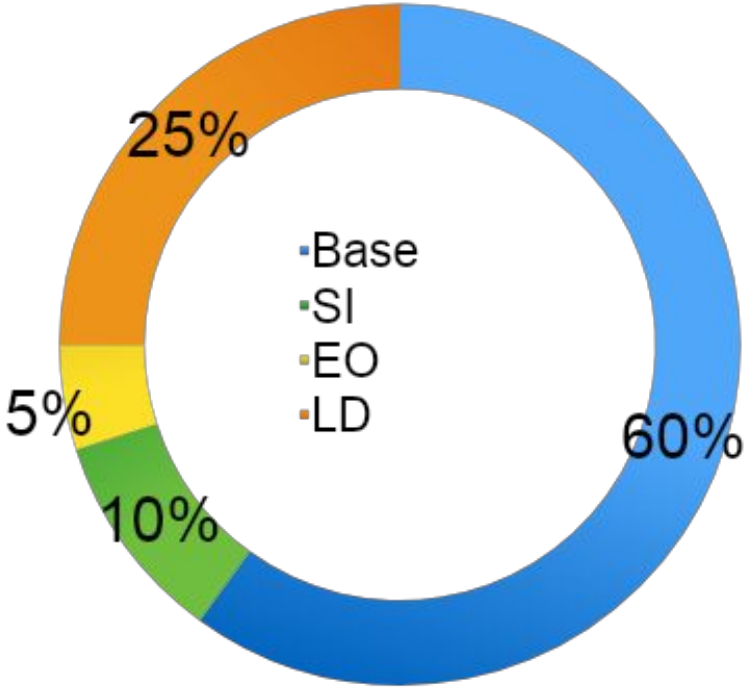
Also offer fundamentally better reliability than QW lasers

QD lasers achieve additional power savings at high temperature.
These savings apply to any optics implementation (DSP, LD, CPO)

Silicon Interposer Optics (“Beyond CPO”)

1. **Ultra low-power silicon interposer electrical interface**
   This lowers electrical signaling power beyond today’s CPO
2. **Potential electrical power savings vs LR: 2 pJ/Bit**
   About 10% compared to 102.4T switch baseline with DSP optics
3. **Requires a very special low-power high-speed SERDES**
   Which cannot drive anything else except this interface

Adding Up The Switch Power Savings



Pluggable SiPh LD optics modules can achieve **25% power savings**

Pluggable best in class LD optics can achieve 3**0% power savings**

25% Linear-Drive (no DSP)
5% Energy efficient modulators
10% Silicon interposer interface
60% Rest of switch power

Thank You

Home

OPEN Compute Project®

Community-driven hyperscale innovation for all.

Loi Nguyen

EVP/GM, Optical and Copper Connectivity Group
Marvell

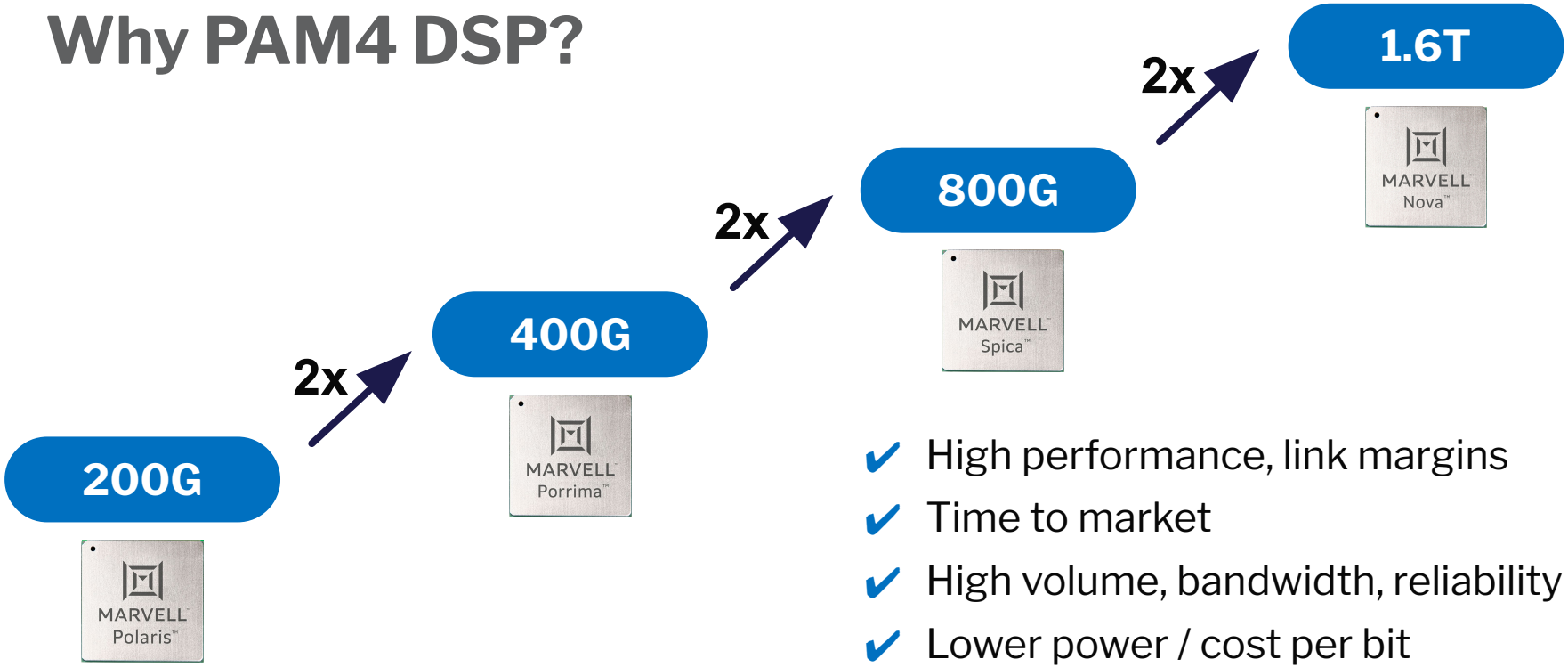
Open Possibilities.

YEARS

OPEN Compute Project®

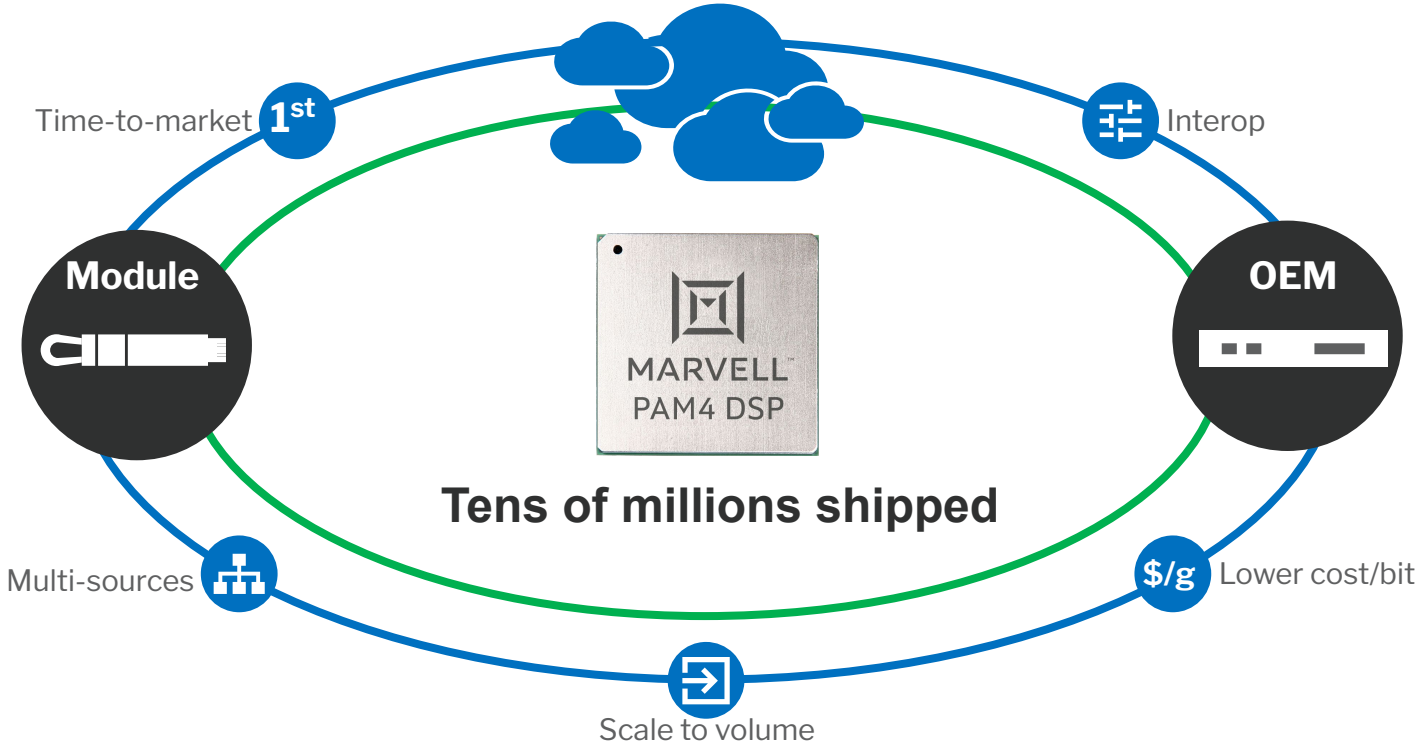
What data center operators want

Why PAM4 DSP?



- ✔ High performance, link margins
- ✔ Time to market
- ✔ High volume, bandwidth, reliability
- ✔ Lower power / cost per bit
- ✔ Open multi-vendor ecosystem

Open multi-vendor ecosystem

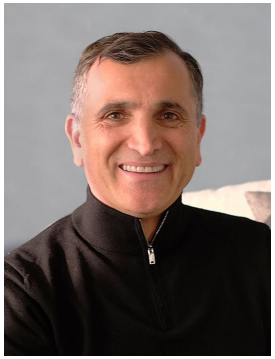


Thank You

Home

OPEN Compute Project®

Community-driven hyperscale innovation for all.



Hamid Arabzadeh

Chairman and CEO, Ranovus

Open Possibilities.

YEARS

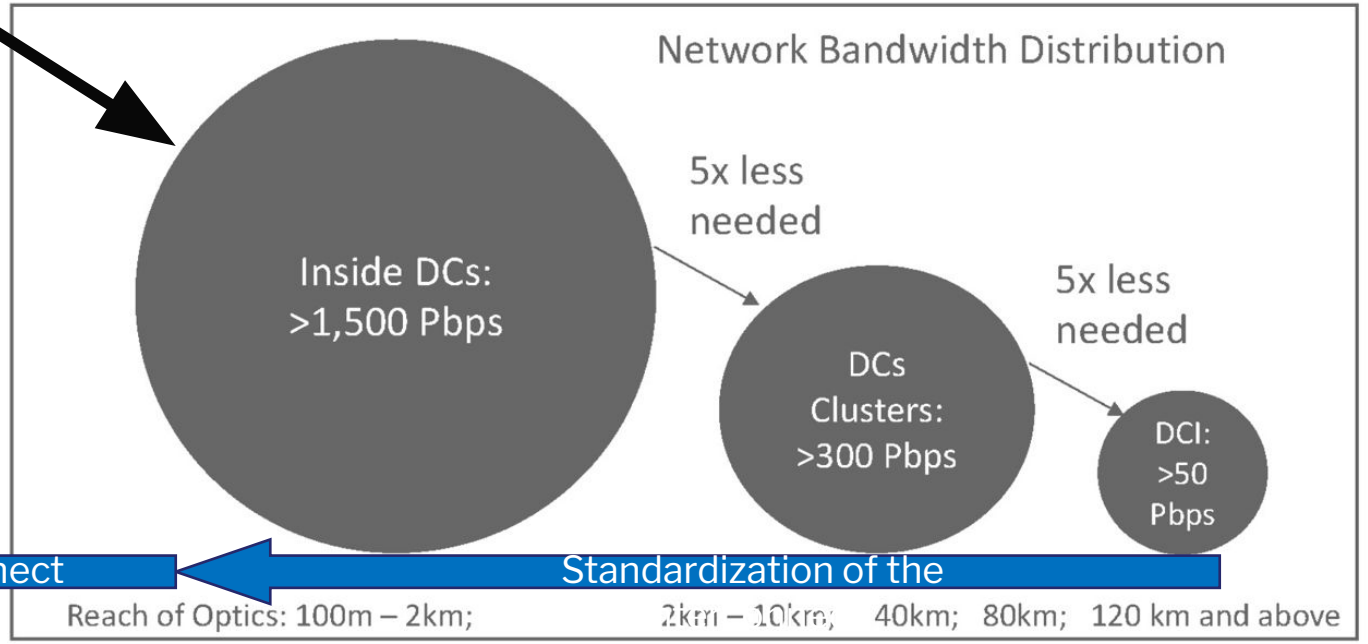
OPEN Compute Project®

Volume of Traffic and Standardization

RANOVUS

AI/ML Interconnect
> !!!!!!! Pbps

Figure 1-3: Bandwidth distribution in networks of Cloud companies

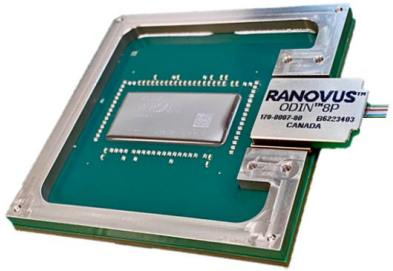


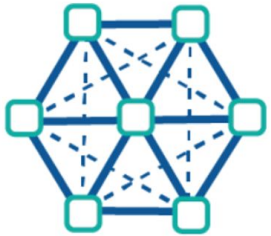
Source: LightCounting

AI accelerator + direct drive 800G CPO interworking with retimed 800G standards-based Ethernet module

RANOVUS™

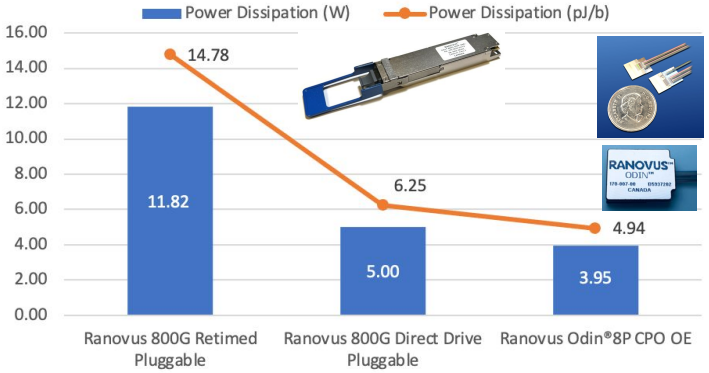
Ranovus' standards compliant Odin® direct-drive CPO 2.0 optical interconnect, with 5pJ/bit energy efficiency, is best-in-class for co-packaged, near-packaged optics and pluggable module form factors





AMD Versal ACAP co-packaged with Ranovus' Odin® direct-drive CPO 2.0

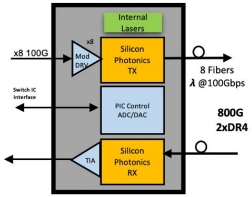
3rd party 800Gbps DR8+ pluggable module and Ranovus 800Gbps DR8+ pluggable module



Device	Power Dissipation (W)	Power Dissipation (pJ/b)	% of Power Dissipation
800G Odin®8P EPIC	2.75		
Integrated Lasers	1.20		
800G Odin®8P OE + ILS	3.95	4.94	33%
800G PAM4 IC	6.4	8.00	54%
Pwr Conv. + Misc	1.47	1.84	12%
OSFP Module	11.82	14.78	100%

All power values are Typical measured numbers

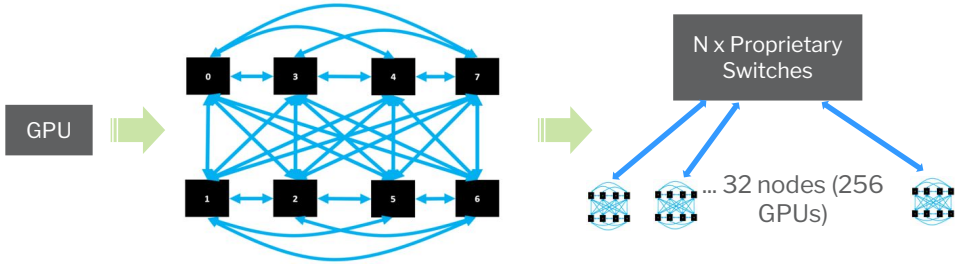
Odin®8P CPO power consumption does not include microcontroller and power conversion



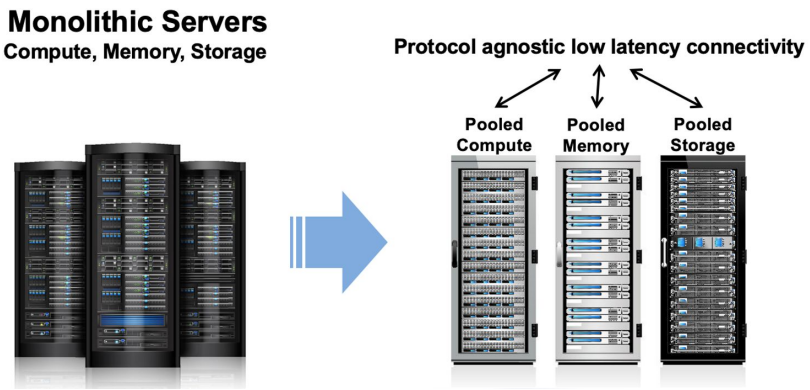
Do AI systems need **retimed** standardized I/O to scale?

RANOVUS

GPU scale-out for AI training workloads



Disaggregated Memory for AI inference workloads



Standardized Ser/Des + Direct Drive optical I/O hold a promising future for AI systems interconnect

Thank You

Home

OPEN Compute Project®

Community-driven hyperscale innovation for all.