Energy Efficient Interfaces

OIF Physical & Link Layer Interoperability Demo
at OFC 2024

Energy Efficient Interfaces @ OFC 2024

- Energy Efficient Interfaces (EEI)

- EEI Interoperability agreements
  - Co-Packaging Framework Document
  - 3.2T Optical Module for Co-Packaging Project
  - ELSFP Project
  - Electrical Interfaces for Co-Packaging

- Interoperability Demonstrations

Copyright @ 2024 OIF

The Challenge!

AI training utilizes large quantities of matrix multiplication

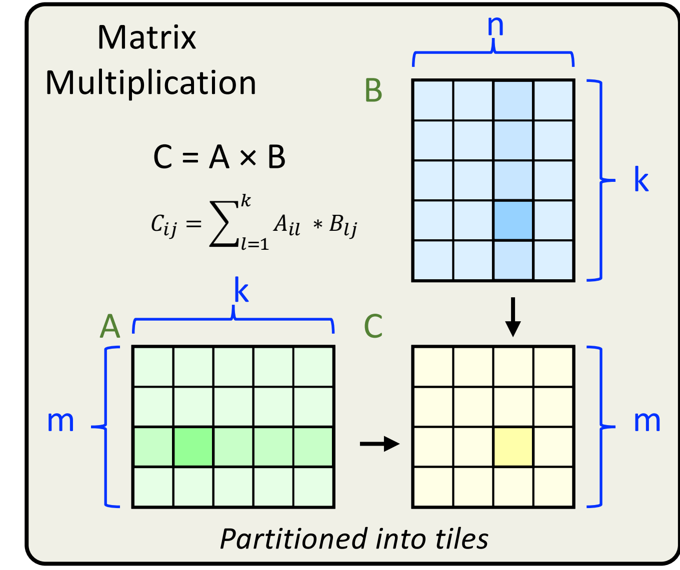
- GPUs are designed to accelerate “multiply and add” operations used in AI matrix multiplication
- Each row in matrix A is paired with every column in matrix B – Lots of computation with lots of parameters!

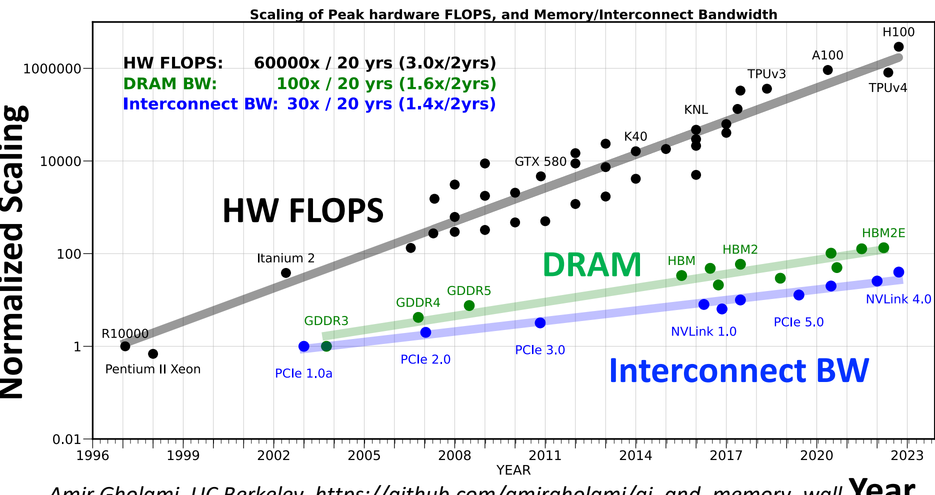
Large AI models can partition the computation into smaller chunks

- Tile computations can be handed off to clusters of local and remote compute accelerators
- However, the completion of a tile in matrix C must wait for all contributing tiles to complete

The time to complete the computation depends on:

- Computation speed
- Interconnect bandwidth
- Memory speed





Year

AI scale-up will drive adoption of optical chiplets to achieve lower latency, energy efficient, & cost-effective interconnections to support large AI models

Copyright @ 2024 OIF

What approaches can we use?

What is needed?

- Larger clusters interconnected with short links
- Energy efficient, high-speed, low latency, dense interconnects that can scale

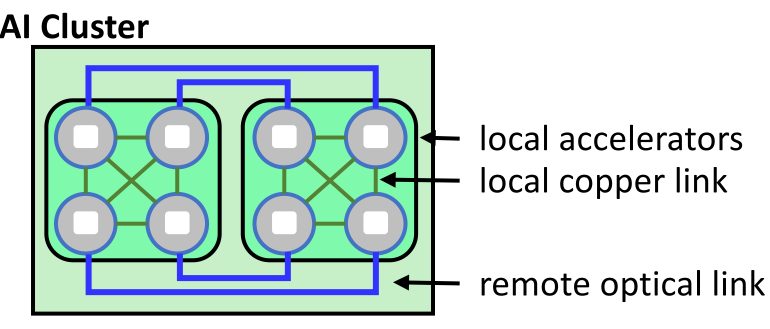
Copper Links

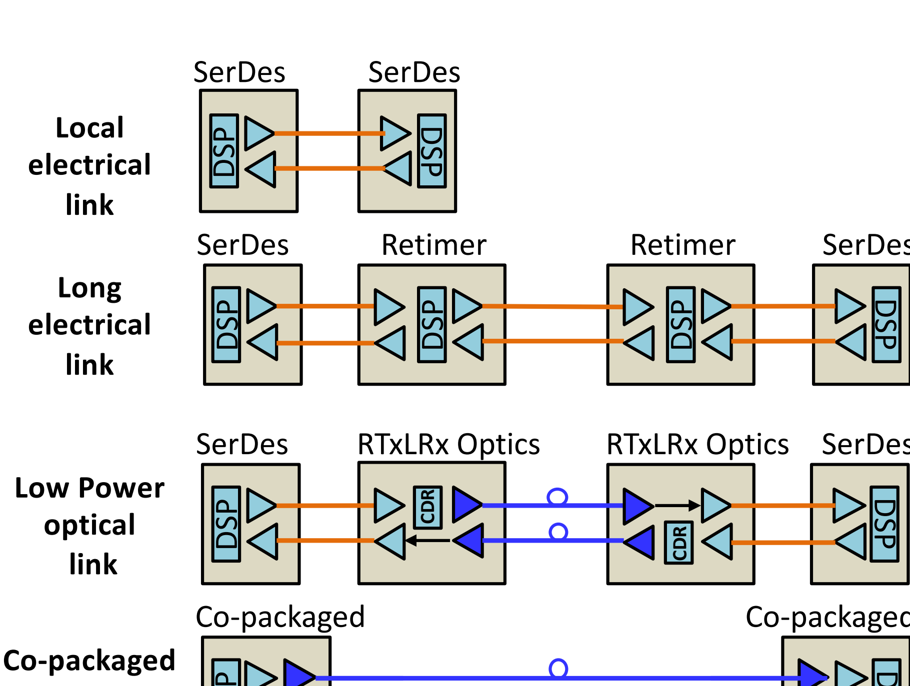
- Copper is ideal for local connections
- As the data rates increase, pure electrical link reach becomes shorter
- Reach can be extended with additional DSP capability and/or with the addition of retimers, trading off additional latency and power

Optical Links

- With some addition of E-O power, the electrical signaling can be converted to optical and then travel without needing additional retimers to restore the signal
- If the electro-optical conversion is co-packaged with the ASIC, additional power and latency can be saved

AI scale-up will drive adoption of optical chiplets to achieve lower latency, energy efficient, & cost-effective interconnections to support large AI models





Copyright @ 2024 OIF

Applications for energy efficient links for AI

Categorization of data center links

Back-end Compute

Mem / CPU / NIC / HBM / Switch / Acc

PCIe/NVLink (low latency)

UEC/InfiniBand (low latency)

Ethernet

Front-end Ethernet Switching

Back-end Switch

Front-end Compute

1, 2, 3, 4, 5

The OIF worked with end-users to summarize their application requirements

BACKEND COMPUTE				FRONT END
Compute type I/Fs *(PCIe/CXL like interfaces)*			Network Type I/Fs	Network Type I/Fs
Compute to I/O	Compute to Memory	Compute to Local Compute	Compute to Remote Compute	Cloud DC
PCIe I/O	Disaggregated memory	Inter-connecting local clusters of accelerators	Inter-connecting remote clusters of accelerators	*Front-End w/ TCP/IP, UDP, etc.* Ethernet (200GbE)
1	2	3	4	5
PCIe	HBM DDR / LPDDR CXL.mem	UPI xGMI NVLink CXL.cache	InfiniBand Ethernet UEC	Ethernet (200GbE)

Backend compute requirements (1–4):

- Low Latency (< 5ns + TOF, EOE & E2E)
- Energy Efficient (< 5 pJ/b)
- Dense Connectivity (1 → 2 T/mm)
- High Degree of Link Accountability

Front end requirements (5):

- Standards based
- Energy efficient
- Potentially high density for some

Adapted from oif2023.269, Ram Huggahalli, Microsoft

Copyright @ 2024 OIF

What kinds of Energy Efficient links are needed?

- Energy efficient, high-speed, low latency interconnects
  - PCIE-like
    - Memory
    - Local Accelerator interconnect
    - I/O
  - Low latency Ethernet
    - Remote accelerator interconnect
  - Ethernet

Back-end Compute

Mem

NIC

CPU

HBM HBM

ACC ACC

Switch

NIC

ACC ACC

HBM HBM

PCIe/NVLink (low latency)

Ethernet

Front-end Ethernet Switch

UEC/Infiniband (low latency)

Back-end Switch

Mem

NIC

CPU

HBM HBM

ACC ACC

Switch

NIC

ACC ACC

HBM HBM

PCIe/NVLink (low latency)

Front-end Compute

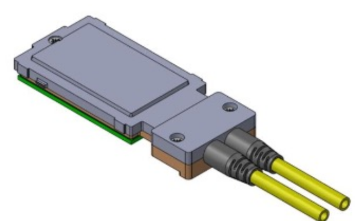
Back-end compute can leverage co-packaged solutions using optical chiplets for dense, energy efficient, low latency interconnections

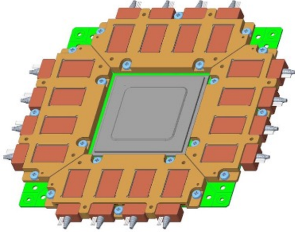
What is the OIF doing?

OIF's Co-Packaging Projects

✓ **Co-packaging Framework Project**

OIF-Co-Packaging-FD-01.0 – Co-Packaging Framework Document



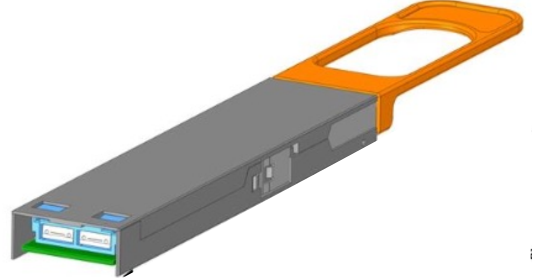


✓ **3.2T Co-packaged Optical Engine**

OIF-Co-Packaging-3.2T-Module-01.0 – Implementation Agreement for a 3.2Tb/s Co-Packaged (CPO) Module

✓ **External Laser Source**

External Laser Small Form Factor Pluggable (ELSFP) Implementation Agreement (August 2023)



Management Interfaces for Optical Engines

- Work in progress

Energy Efficient Interfaces for AI

System Vendor Requirements Document for Energy Efficient Interfaces

- Document the EEI requirements as provided by the end-users for AI/ML optical and electrical links

Energy Efficient Interface Framework

- Study and initiate new standards for dense, low power, low latency links for AI/ML

RTxLRx Project

- Develop a standard around a partially retimed Ethernet optical engine for lower latency and low power applications
- (Retimed Transmitter, Linear Receiver)

Copyright @ 2024 OIF

Energy Efficient Interfaces @ OFC 2024

- Energy Efficient Interfaces (EEI)
- EEI Interoperability agreements
  - Co-Packaging Framework Document
  - 3.2T Optical Module for Co-Packaging Project
  - ELSFP Project
  - Electrical Interfaces for Co-Packaging
- Interoperability Demonstrations

Copyright @ 2024 OIF

Co-Packaging Application Spaces

Framework Project

Data-Center Networking | AI Training / ML | Disaggregation

End Node

Sparse →

Optical Engine

Switching Node

Dense →

Electrical Engine

Low Latency I/F

Ethernet I/F

Compute End-Point

DRAM
Server Chip
PCIe
NIC
CPO

CPO CPE
HBM NPU/GPU HBM
CPO CPE

Server CPU Package
Local Memory
Server Chip
Server Chip
Local Memory
CPO
CPO
Compute Fabric

Switches

Ethernet Switch ASIC
CPO

High Perf. Switch ASIC

Compute Fabric
CPO CPO CPO CPO
Switch
NIC MC MC
CPO
Ethernet
Remote Memory

100 – 800Gb/s	Bandwidth	High (Multi-Tbps per compute node)
µs - ms	Latency	Low (<200ns RT memory latency)
10 – 50 FITs	Reliability	Very High (< 1 FIT)

Source: OIF CPO Webinar, Ram Huggahalli 10-14-2020

Copyright @ 2024 OIF

Co-Packaging Architectures (1)

Framework Project

Co-Packaged using socket for engine

Copper Cables Or Fibers

OE or EE

Engine Substrate

soldered

ASIC

socketed

Front Panel

Co-Packaged Assembly Substrate

Host PCB

Co-Packaged with soldered engine

Copper Cables Or Fibers

OE or EE

Engine Substrate

soldered

ASIC

soldered

Front Panel

Co-Packaged Assembly Substrate

Host PCB

Copyright @ 2024 OIF

Co-Packaging Architectures (2)

Framework Project

Co-Packaged using copper cable assembly

Copper Cable
or
To Front Panel or Backplane
Copper Cables Or Fibers
soldered
ASIC
OE or EE
Co-Packaged Assembly Substrate
Engine Substrate / RDL
Host PCB
Front Panel

Near-Packaged using socketed engine

Copper Cables Or Fibers
ASIC
OE or EE
socketed
ASIC Substrate / RDL
Engine Substrate
socketed
Co-Packaged Assembly Substrate
Host PCB
Front Panel

Copyright @ 2024 OIF

Interfaces Studied for Interoperability

Framework Project

Thermal, Mechanical & Environmental

Copper Cables Or Fibers

Electrical

OE or EE

ASIC

soldered

Engine Substrate

soldered

Co-Packaged Assembly Substrate

MIS

Power

Socket

Optical (Electrical) & RLS

Front Panel

Host PCB

Reliability and Repairability

Application Example

- **Switch Generation:** 51.2Tb/s
- **Lane Speed:** 106 Gb/s
- **Interface Architecture:** XSR based AUI, 400G-FR4 PMD
- **Motivation:** System power reduction, ecosystem & operational readiness

Copyright @ 2024 OIF

Energy Efficient Interfaces @ OFC 2024

- Energy Efficient Interfaces (EEI)

- EEI Interoperability agreements
  - Co-Packaging Framework Document
  - 3.2T Optical Module for Co-Packaging Project
  - ELSFP Project
  - Electrical Interfaces for Co-Packaging

- Interoperability Demonstrations

Copyright @ 2024 OIF

Example System Attachment

3.2T Optical Module

- 16 x 3.2T Modules = 51.2T Switch Capacity

Co-Packaged Assembly Substrate (Interposer)

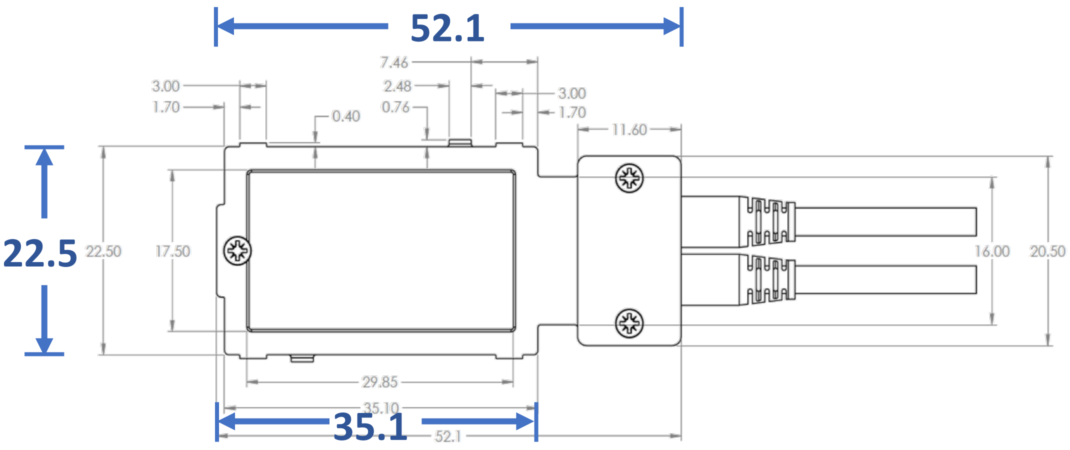
Channel components cross-section

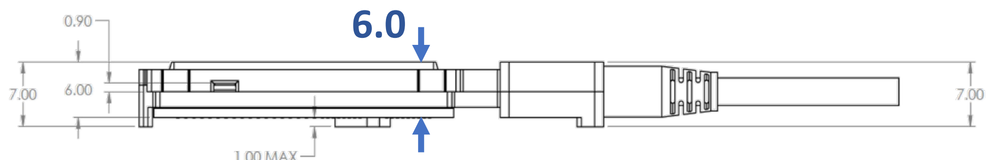
Copyright @ 2024 OIF

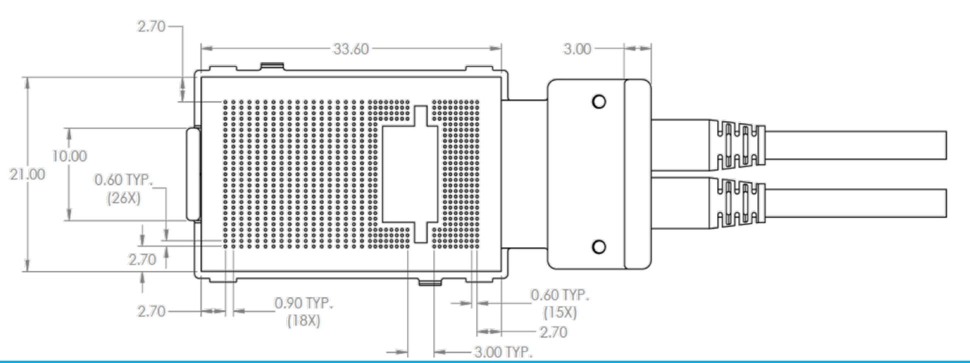
3.2T Module Dimensions

3.2T Optical Module

- 32 x 112G XSR to Standard Optics:
  - 8 x 400G DR4
  - 8 x 400G FR4 (incl. 200G mode)
- Copper Cable Assembly compatible
- Power capability:
  - 56W (Internal Laser option)
  - 48W (External Laser option)







Copyright @ 2024 OIF

LGA Pin Map

3.2T Optical Module

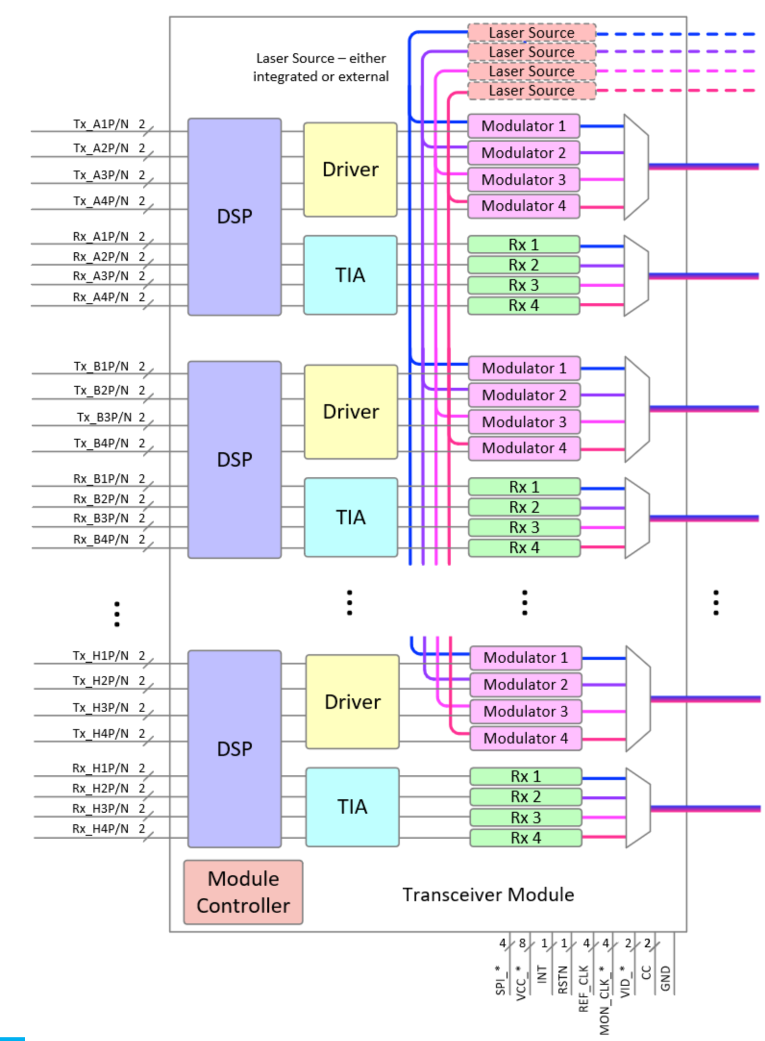
- Supply rails: 12V, 3.3V, 2.6V, 1.8V, 1.2V, 0.9V, 0.7V
- Comms Electrical: 1.2V SPI
- Comms protocol: CMIS
- 400G and 800G (2x400G) port grouping defined
  - *For low power modes and 2x400G-FR4 cable assignment*

Copyright @ 2024 OIF

3.2T Optical Module Functionality

3.2T Optical Module

- FR Module example ->
- *How does this all fit in?*
  - *3D integration*
  - *Die/functionality integration*
    - *Optics (Laser + Modulator + PD)*
    - *EIC (Driver/TIA/Control)*



Copyright @ 2024 OIF

Energy Efficient Interfaces @ OFC 2024

- Energy Efficient Interfaces (EEI)

- EEI Interoperability agreements
  - Co-Packaging Framework Document
  - 3.2T Optical Module for Co-Packaging Project
  - ELSFP Project
  - Electrical Interfaces for Co-Packaging

- Interoperability Demonstrations

Copyright @ 2024 OIF

Why ELSFP?

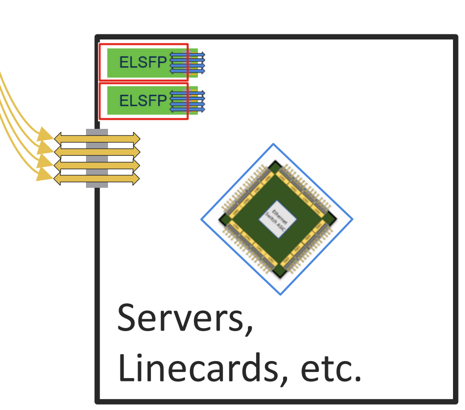
ELSFP Project

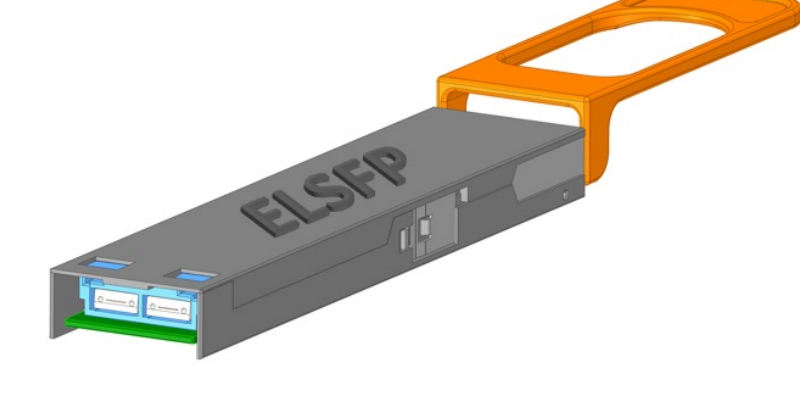
- OIF defining common External Laser Pluggable
- Industry need for co-packaged and near-packaged systems
  - Systems need faceplate density
  - External laser modules need common specification for economies of scale
- Form factor to span multiple system generations
  - Plan for optical & thermal scaling

External Laser Small Form Factor Pluggable (ELSFP)

Examples: Switches and Special Purpose ASICs

- ELSFPs provide CW laser power for optical engines (OEs).
- Decreases thermal power density in the system
- Each large system will likely need multiple (i.e. 8 or 16) ELSFPs
- The light from a given ELSFP can feed more than a single OE.
- A pluggable form factor helps to ensure total system reliability and a “hot swap” replacement if a single laser or ELSFP module fails.
- Eye safety is achieved by a blind mate optical connector internal to the system.





Copyright @ 2024 OIF

Initial Technical Concept

ELSFP Project

Density

- Blind mate pluggable
- Width similar to OSFP (16 modules wide with standard management I/O)

Commonality

- Industry standard 3.3V Supply
- CMIS (Common Management Interface Specification)

Scaling

- Optical Power Classes
- Thermal Power Classes
- Belly-to-belly configurations
- Riding heat sink for system flexibility
- 2 “MT like” ferrules for future proofing
  - Support for 8 PM fibers per MT
  - Support for multiple OE modules

Copyright @ 2024 OIF

Single Port ELSFP Design

ELSFP Project

Press fit host connector

Riding heat sink

Optical connectors snap into host connector housing

Separately applied connector and cage

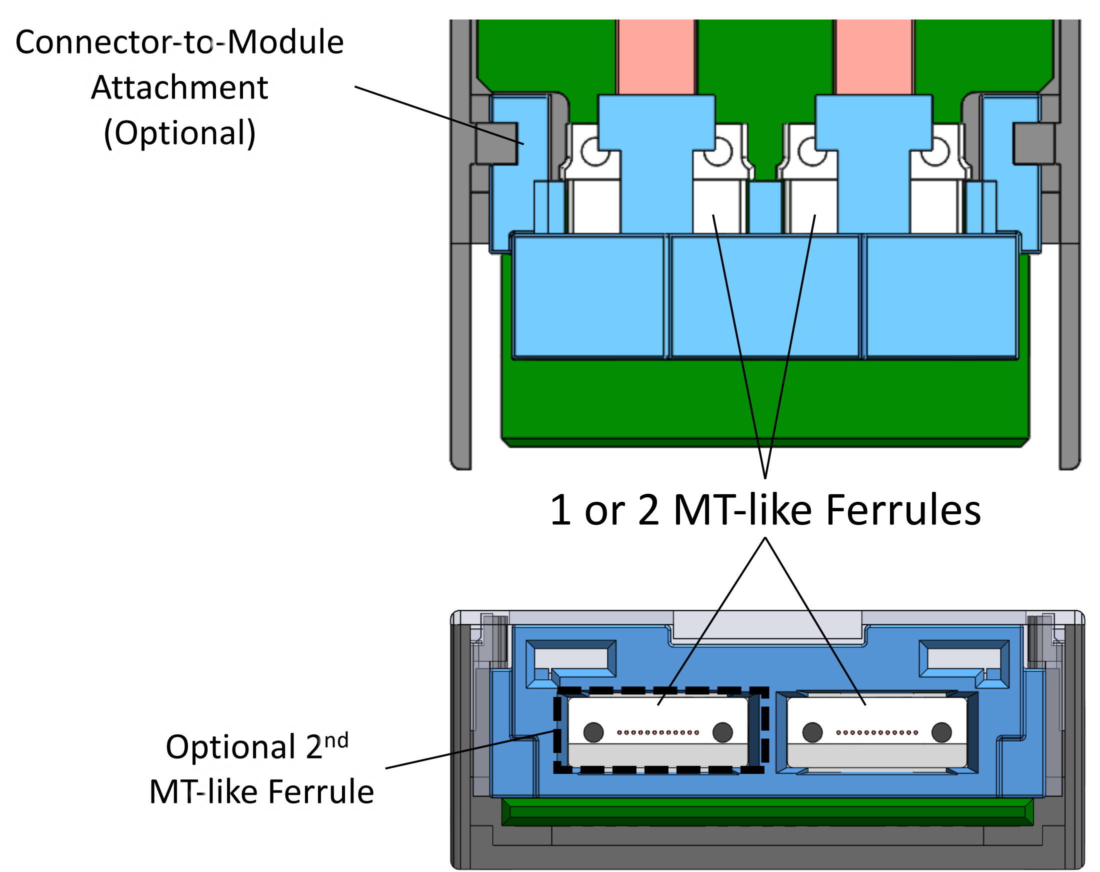
OSFP style latch

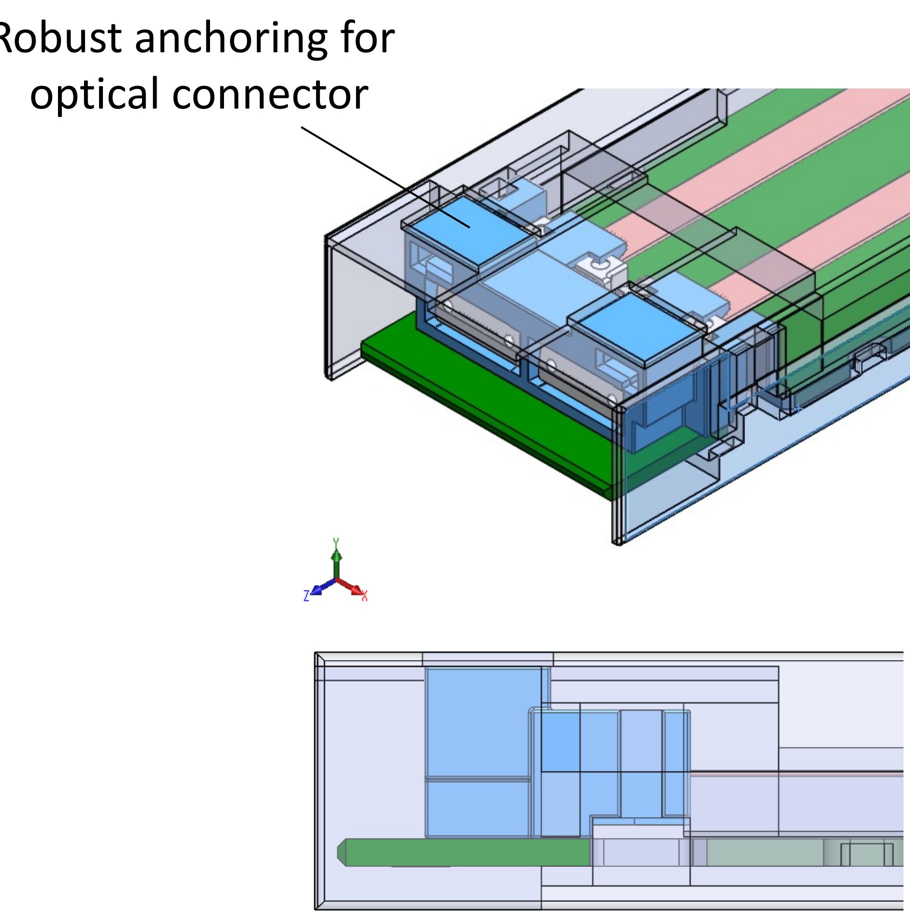
Shroud providing paddle card and optical connector protection

Copyright @ 2024 OIF

ELSFP Module-Side Optical Connector

ELSFP Project

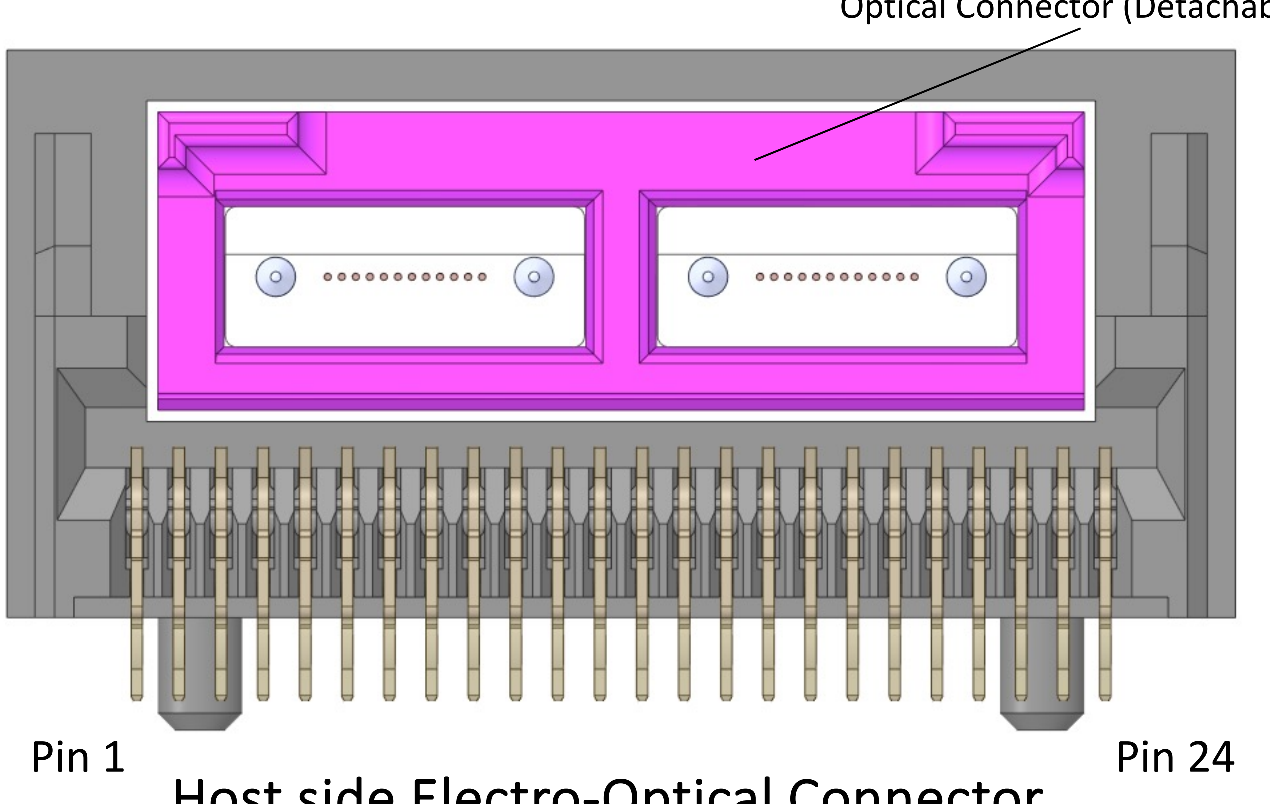




Copyright @ 2024 OIF

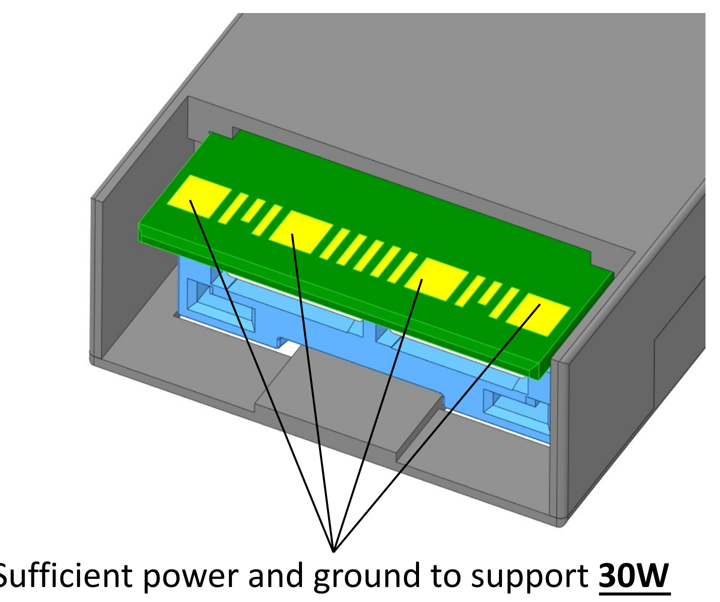
ELSFP Electro-Optical Connector

ELSFP Project



Host side Electro-Optical Connector

Module Bottom side Electrical Contacts



Sufficient power and ground to support **30W**

Additional pins for control/management, laser safety (i.e. presence pin), and spares for future proofing
Optical connector sub-assembly (pink) is separable from the board mounted electrical connector sub assembly

ELSFP Optical Power Classes

ELSFP Project

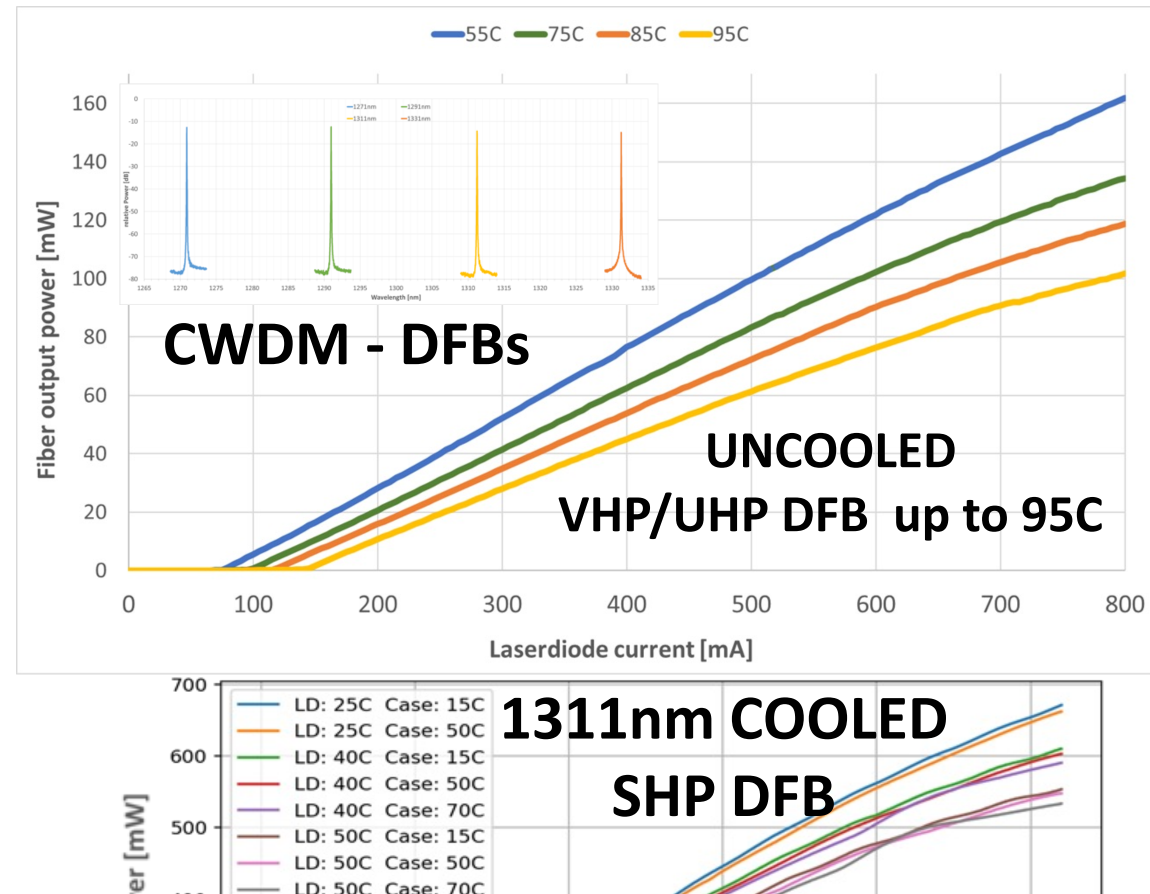
ELSFP Optical Power Classes	Power/λ/Core +/- 1.5dB
Super Low Power - SLP	2dBm
Ultra Low Power - ULP	5dBm
Very Low Power - VLP	8dBm
Low Power - LP	11dBm
Medium Power - MP	14dBm
High Power - HP	17dBm
Very High Power - VHP	20dBm
Ultra High Power - UHP	23dBm
Super High Power - SHP	26dBm

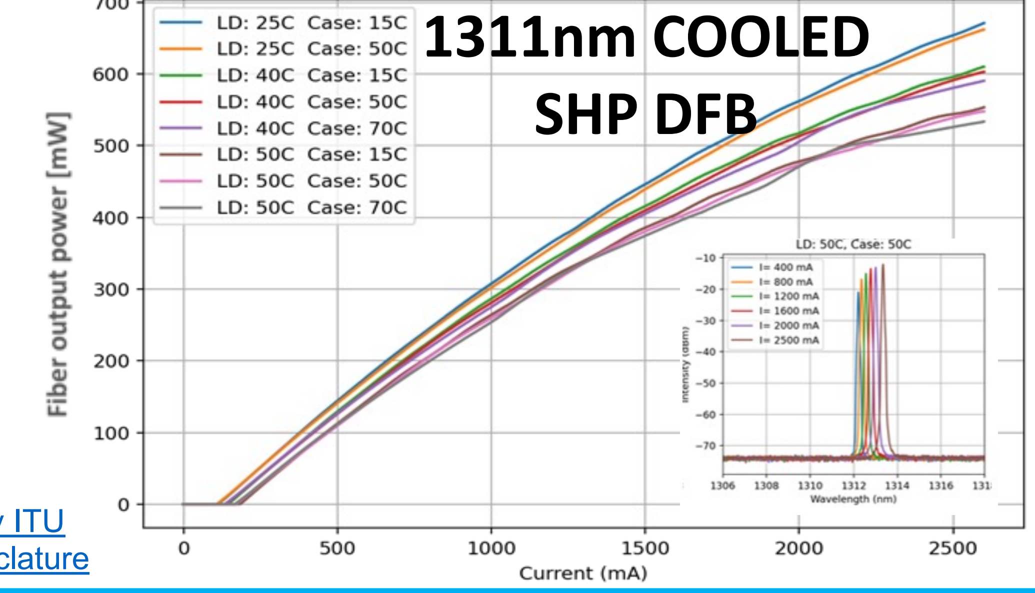
Combs

Single-Channel

Multi-Channel

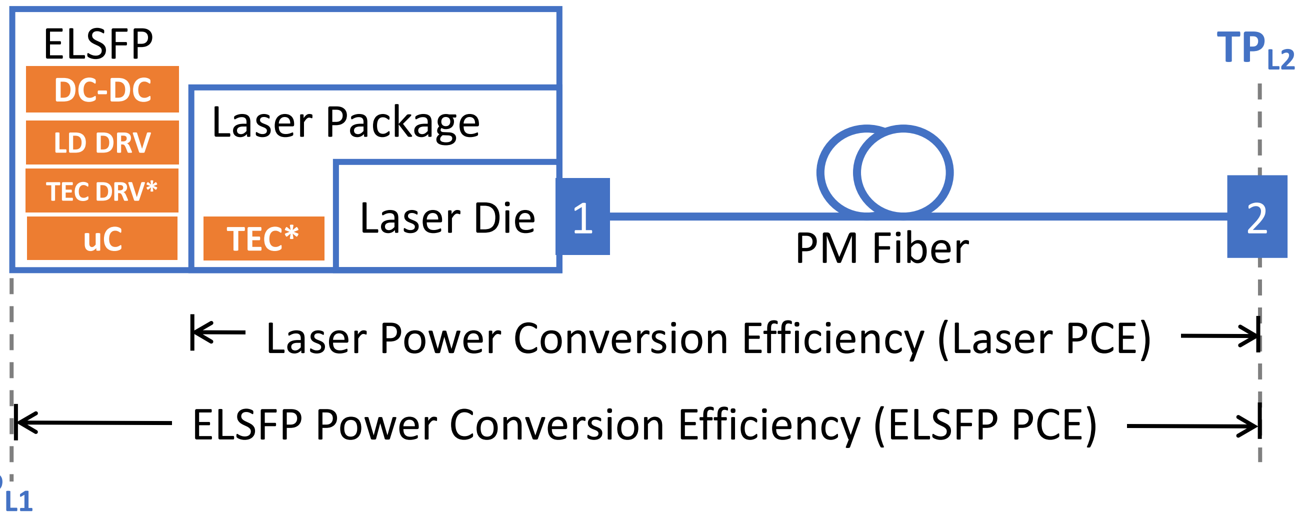
*Naming convention inspired by ITU Radio Frequency Band Nomenclature

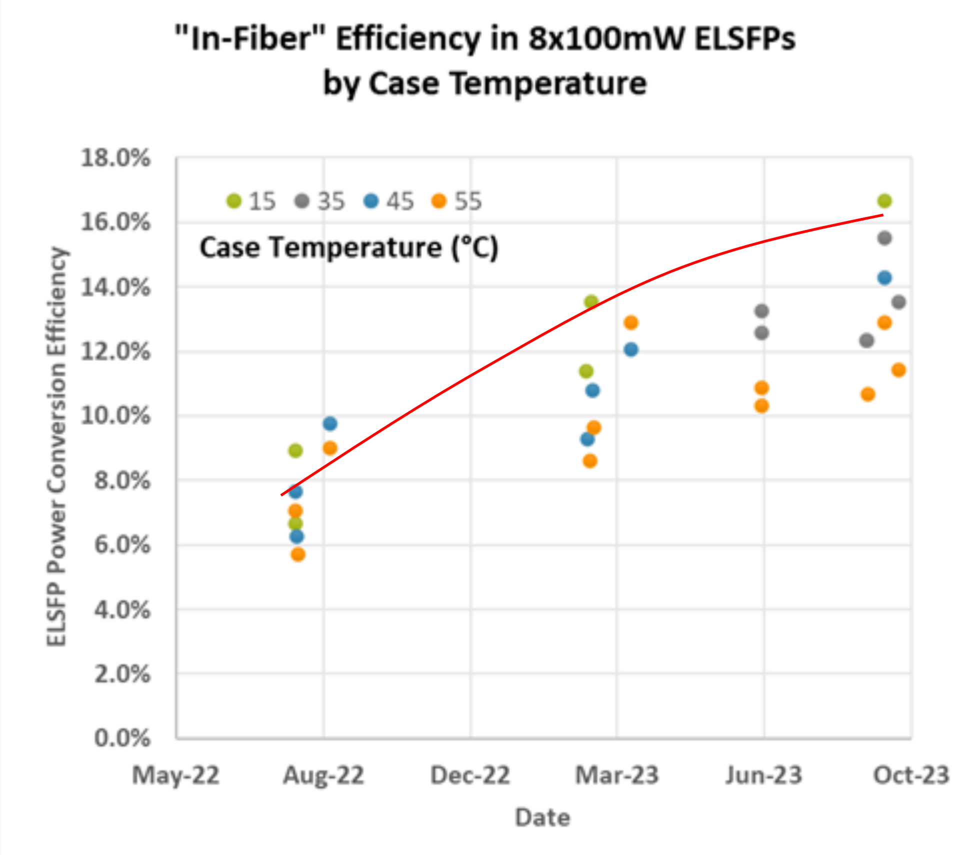




Copyright @ 2024 OIF

ELSFP's eco-system drives innovation





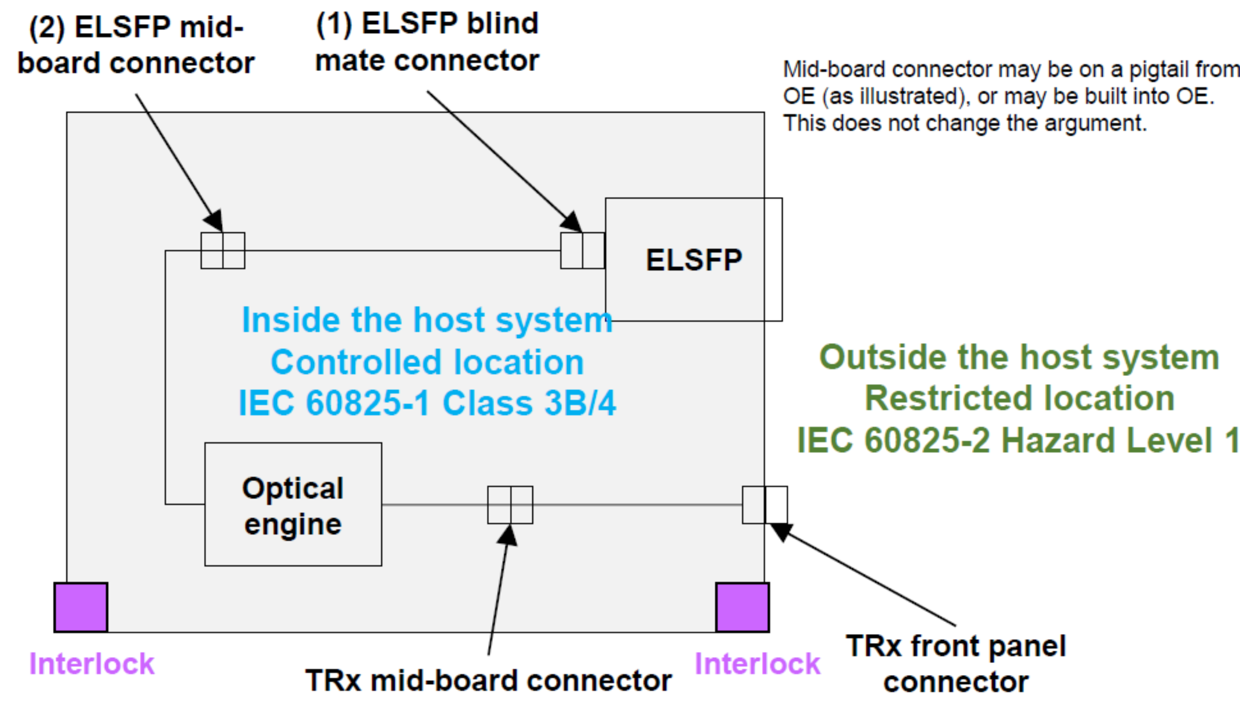
The ELSFP's eco-system continues to innovate and has yielded impressive improvements in energy efficiency (PCE), a key component of next generation energy efficient interfaces

Copyright @ 2024 OIF

Eye Safety

ELSFP Project

ELSFP's blind mate optical connector paired with a system interlock enables a safer co-packaged system implementation for users.



Similar to EDFAs with powerful CW lasers, Class 3B and 4 lasers can be used inside ELSFP and systems can be deployed in unrestricted locations.

Copyright @ 2024 OIF

Energy Efficient Interfaces @ OFC 2024

- Energy Efficient Interfaces (EEI)

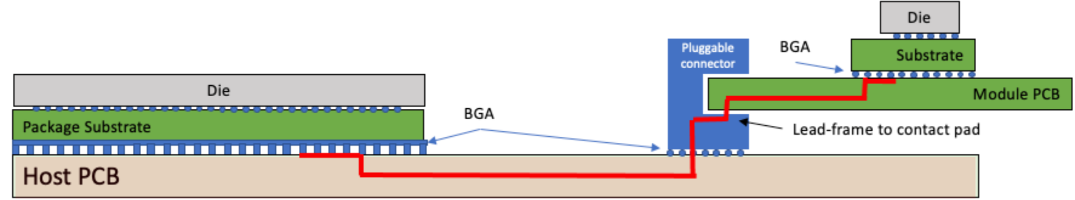
- EEI Interoperability agreements
  - Co-Packaging Framework Document
  - 3.2T Optical Module for Co-Packaging Project
  - ELSFP Project
  - Electrical Interfaces for Co-Packaging

- Interoperability Demonstrations

Copyright @ 2024 OIF

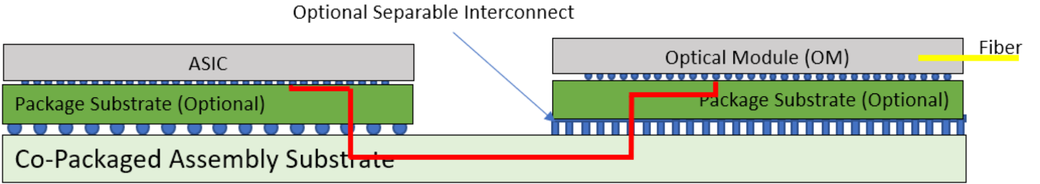
CEI – An Essential Building Block for Co-packaging

Pluggable Module Channel Example Illustration



- Channel loss: 16dB ball to ball (22-24dB bump to bump)
- Typical pluggable connectors: IL of ~1dB with RL of -10dB @26.5GHz

CPO/NPO Channel Example Illustration

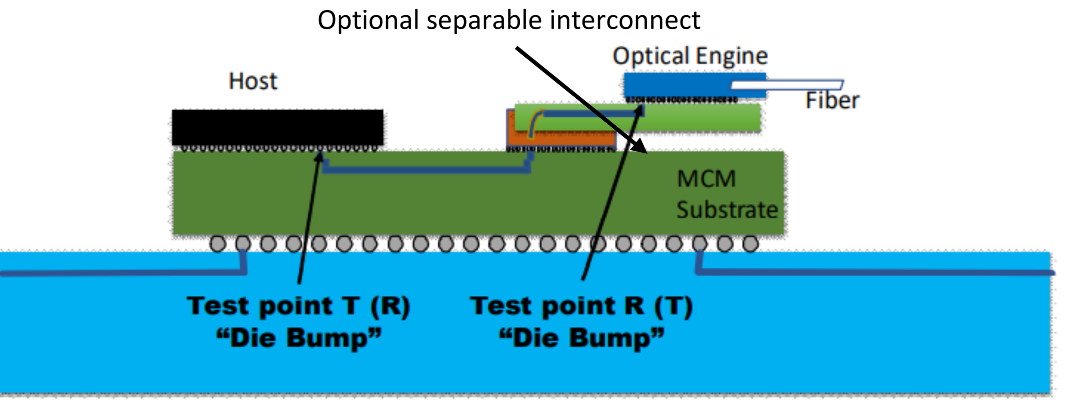


- Channel loss: CPO – 10dB bump to bump; NPO – 13dB bump to bump
- Optional separable interconnect performance example: LGA socket: IL of ~0.05dB with RL of -40dB @26.5GHz (*oif2020.341.01, Nathan Tracy*)
- Avoids/reduces major discontinuities.
- Optical modules are not end user pluggable.

- Significant power saving opportunity over VSR to be captured.
- A broad interoperable ecosystem is the key to success and can only be achieved through standardization.

Copyright @ 2024 OIF

CEI-112G-XSR-PAM4 for Co-packaging



Category	IL at Nyquist (Max, dB)	BER (Max)
CAT1	**10**	**1e-6**
CAT2	10	1e-8
CAT3	8	1e-9

- Baud rates supported: 36 Gsyms/s to 58 Gsyms/s
- Based on loss and jitter budgets between TX and RX using copper signal traces in a SIP(System in a Package) to enable low power consumption
- Three channel categories are defined, allowing optimization for various applications.
- Timeline
  - Project started in April 2018.
  - Draft specification is becoming technically stable. Few pending items to be addressed.

CEI-112G-XSR+ -PAM4 for Near Packaging

- The emergence of Near Package Optics (NPO) Architecture
  - Co-packaging requires significant package substrate size increase and technology advancement, which adds risk to goals of availability, cost and multi-vendor support.
  - Instead of a monolithic package approach, Near Packaging relies on advanced PCB technology for dense high-speed routing without significant power penalty.
  - Near Packaging architecture takes advantage of existing technologies and more robustly enables an open ecosystem implementation.
- Additional margin also strengthens a broader supply base for co-packaging implementation and adoption.

- Baud rates supported: 36 Gsyms/s to 58 Gsyms/s
  - Optimize for Ethernet rate @ 106.25Gbps – the key application for CPO/NPO
  - Insertion loss < 13dB @ 26.5625GHz Nyquist bump to bump with up to 1 separable interconnect.
- Enable the lowest practical energy consumption (pJ/b) implementation.
- Leverage specification methodology and other work from existing CEI 112 projects.

Optional Separable Interconnect

ASIC

Package Substrate (Optional)

Optical Module (OM)

Fiber

Package Substrate (Optional)

Co-Packaged Assembly Substrate
(PCB Interposer)

Copyright @ 2024 OIF

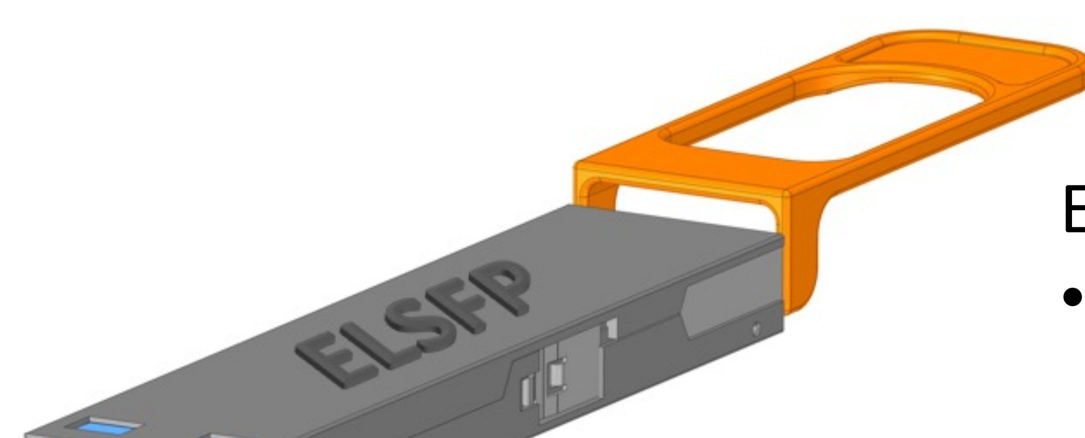
Energy Efficient Interfaces @ OFC 2024

- Energy Efficient Interfaces (EEI)
- EEI Interoperability agreements
  - Co-Packaging Framework Document
  - 3.2T Optical Module for Co-Packaging Project
  - ELSFP Project
  - Electrical Interfaces for Co-Packaging
- Interoperability Demonstrations

Copyright @ 2024 OIF

Interoperability Demonstrations:
External Laser Small Form Factor: ELSFP

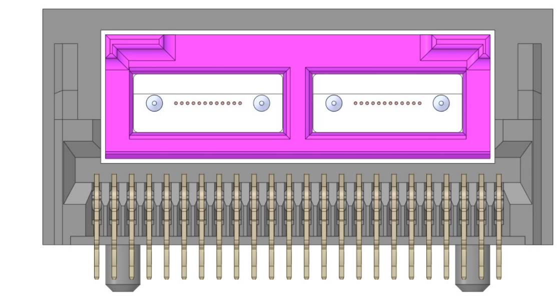
SENKO®
Advanced Components



ELSFP from multiple vendors

- 20-23 dBm output power

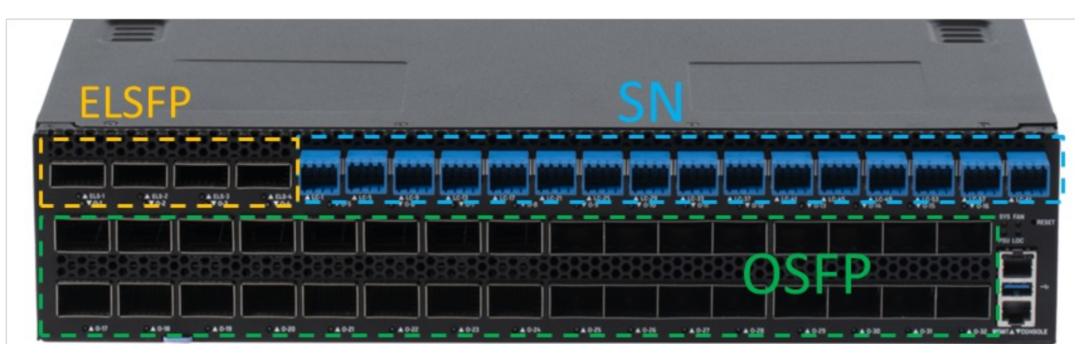
Blind mate connector cleaning system



Copyright @ 2024 OIF

Interoperability Demonstrations: Active NPO Switch

- NPO Switch with both electrical and optical interfaces
- OIF's 3.2T OE and LPO optical interfaces
- External laser sources (ELSFP)
- Remote optical link



Copyright @ 2024 OIF

Interoperability Demonstrations:

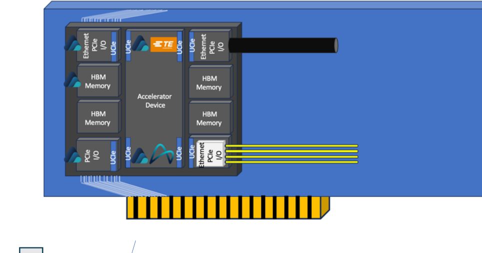
Conceptual EEI (Energy Efficient Interfaces)

Showing examples of both Near-term and Long-term architectures

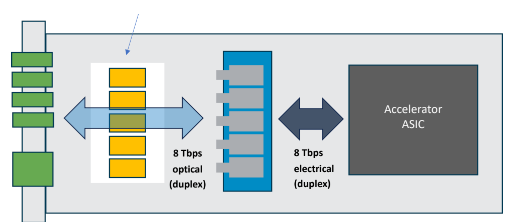
Front-end Ethernet switch

NIC
NIC
Integrated FE NIC
PCIe Bus
CPU
MCM
NIC
HBM
HBM
PCIe Switch
NIC
ELSFP

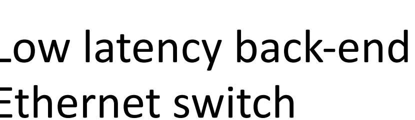
Next-gen Integration



Next-gen co-packaged accelerator

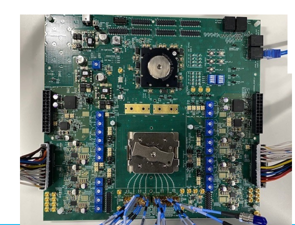


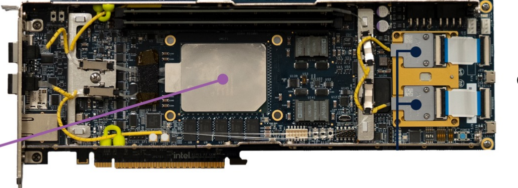
8T accelerator



Low latency back-end Ethernet switch

Optical PCIe low power interface





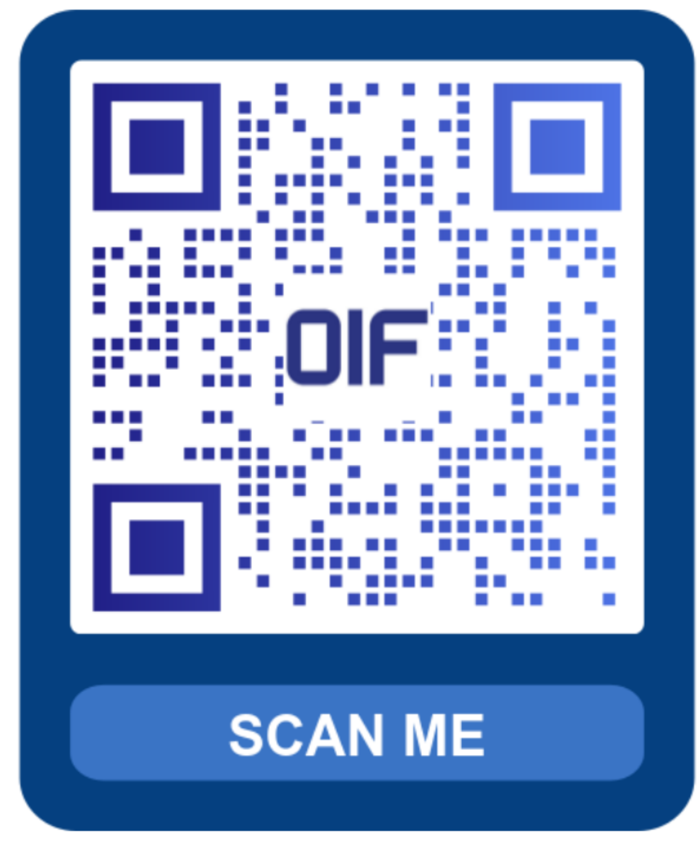
8T accelerator with integrated XPU

USCONEC

wistron

Amphenol

OIF
PLL INTEROP DEMO
OFC 2024


OIF
SCAN ME