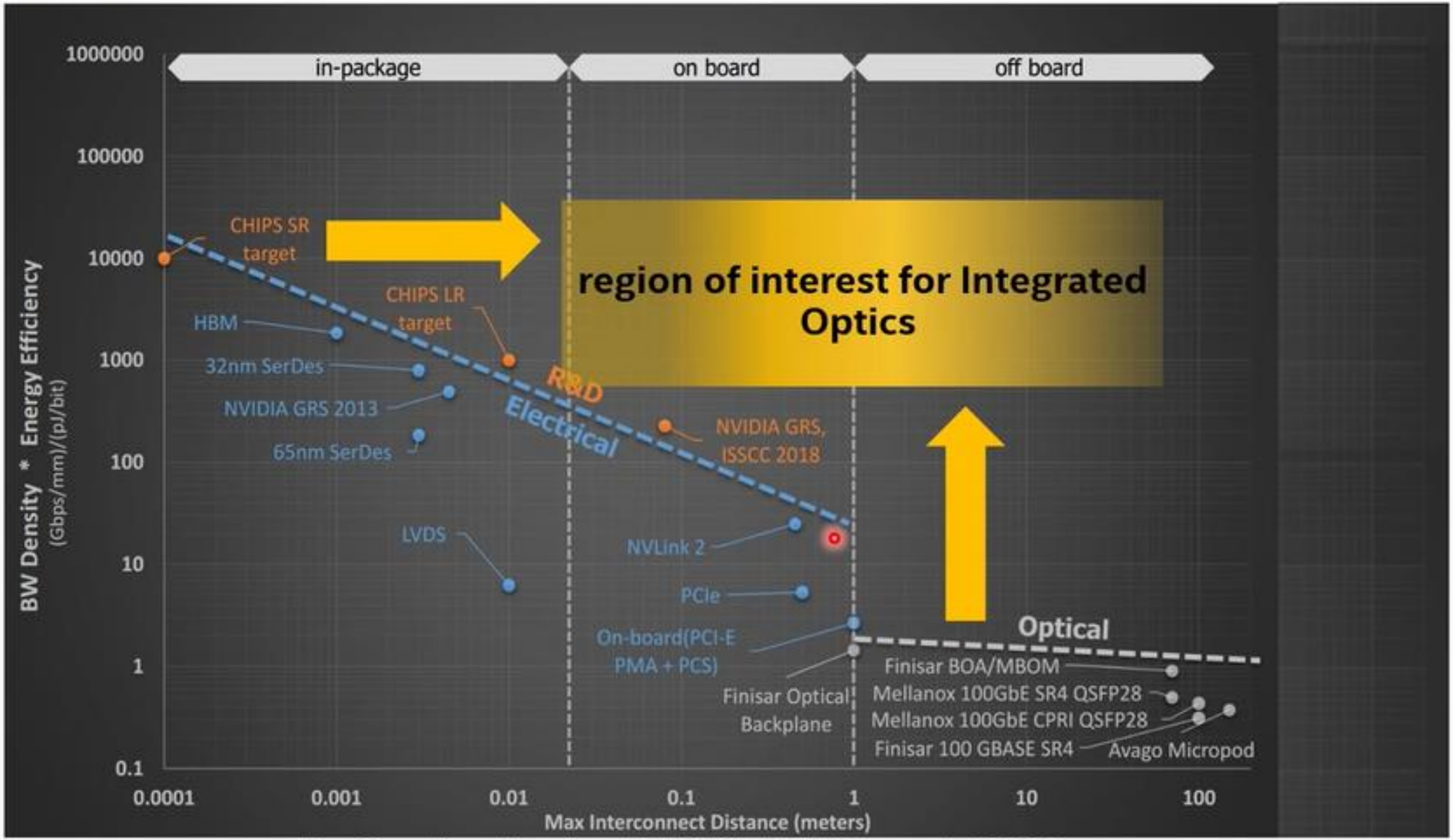
Optical Transceiver Overview & Optical Testing

Yi-Shing Chang
Principal Engineer, Intel Corp
Adjunct Professor, GSAT, NTU

Outline

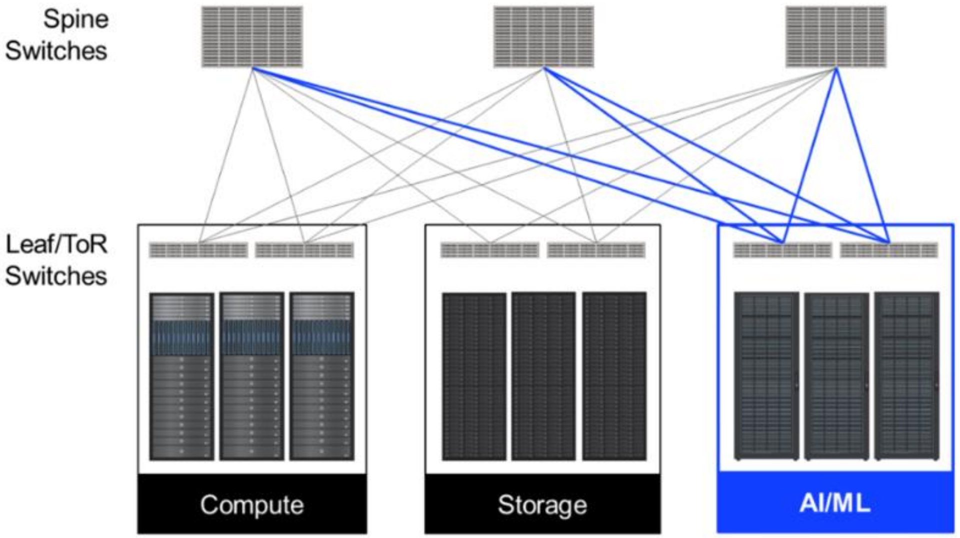
- Motivation: the era of integrated optics
- Optical Transceiver Components Overview
- 100G vs 400G & above
- Transceiver Module Test Overview
- AI/ML Applications



Dr. Gordon Keeler, DARPA ERI Summit 2019

AI Era – Datacom Opportunity

Time to reach **one million users**	
ChatGPT	**5 days**
	2.5 months
facebook	10 months
	2 years
NETFLIX	3.5 years

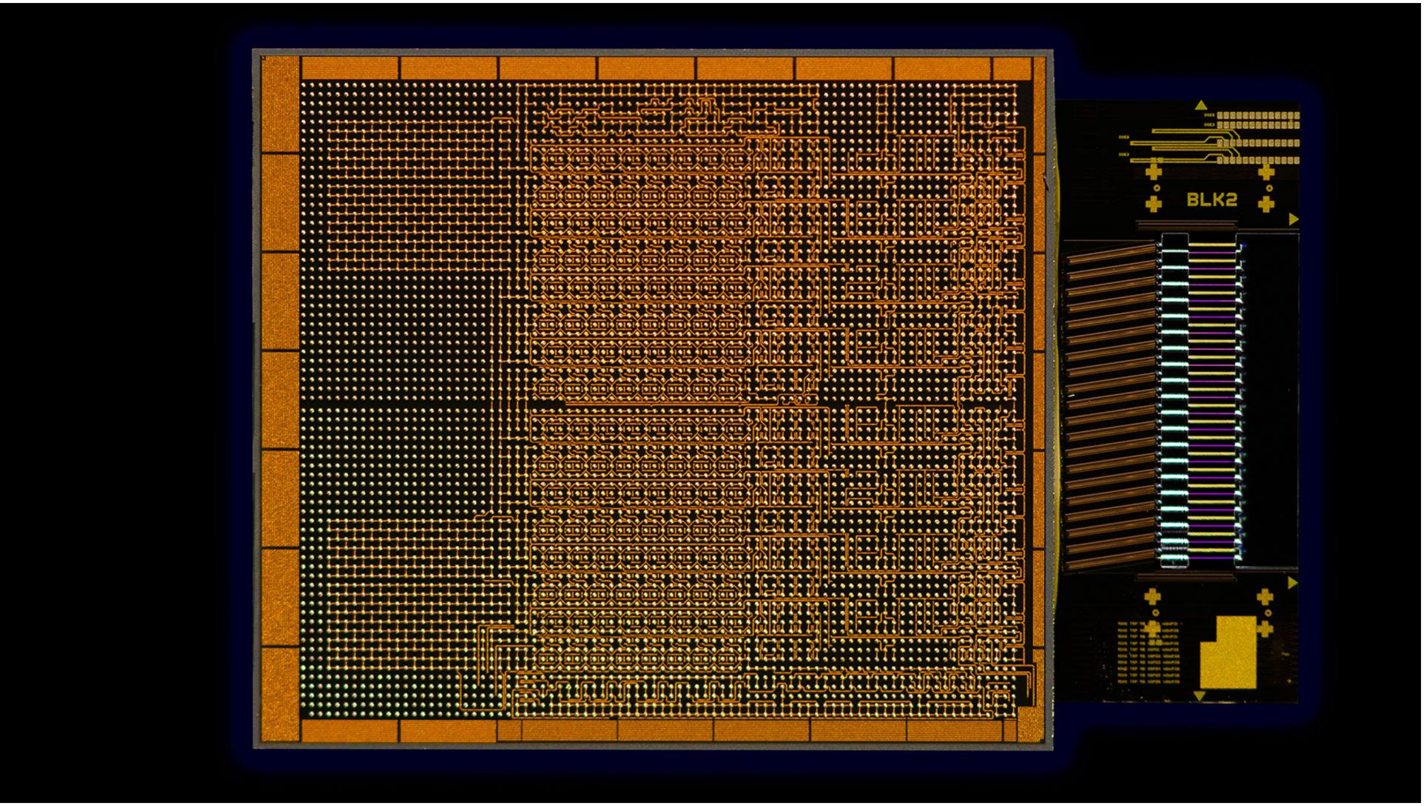


AI/ML servers and their connectivity fabric, alongside connectivity for traditional compute and storage servers, increase the number of optical links in data centers

A new accelerated compute portion of the network (Level 0 or back-end), which consists of AI/ML servers and accelerated compute devices, bolts onto the traditional network alongside traditional compute and storage. Optical interconnects are used at all levels in this network

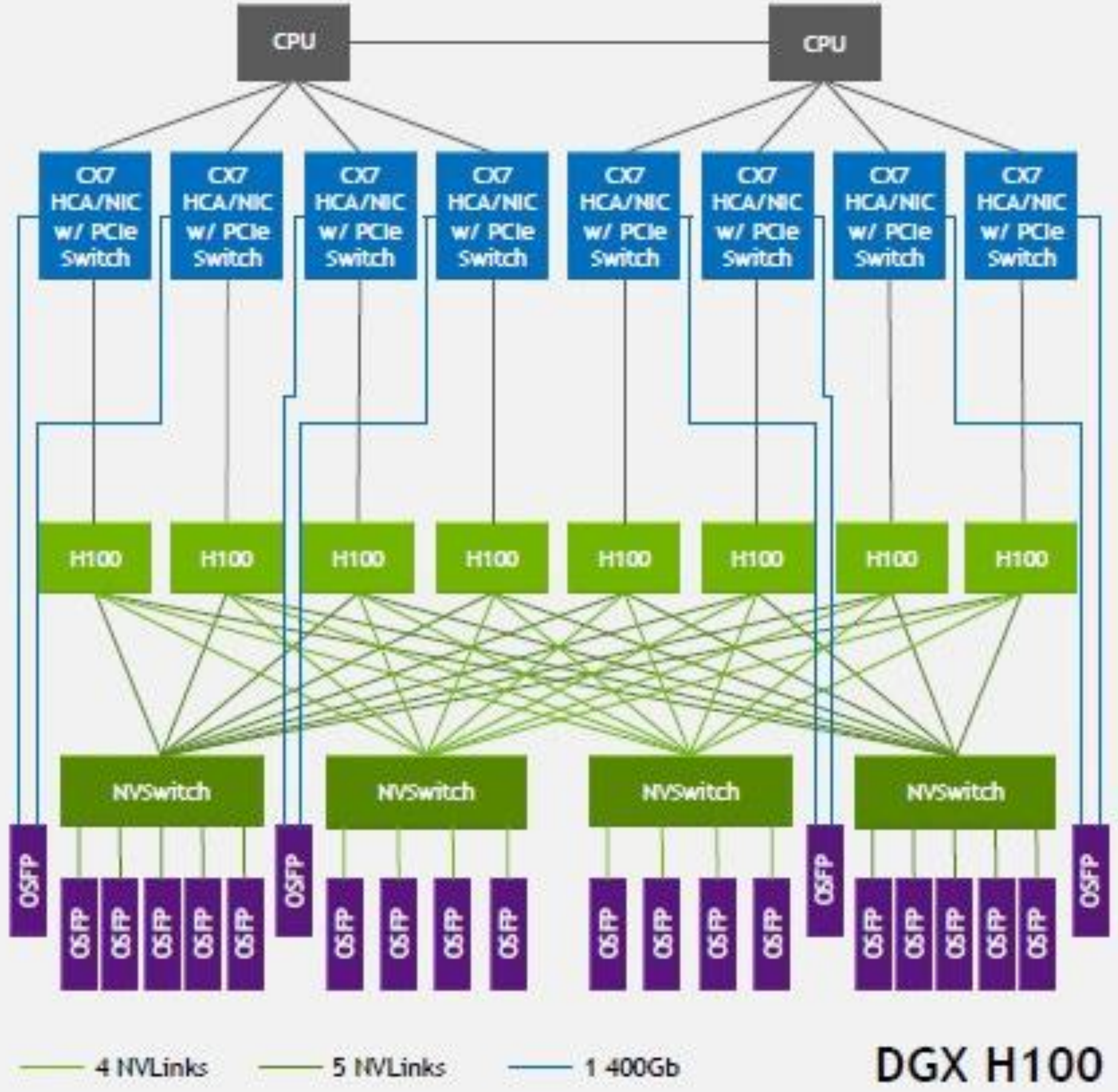
https://www.coherent.com/news/blog/enabling-ai-next-gen-datacenters

Optical Compute Internet Chiplet



https://www.intel.com/content/www/us/en/newsroom/news/intel-unveils-first-integrated-optical-io-chiplet.html

- This first OCI chiplet is designed to support 64 channels of 32 gigabits per second (Gbps) data transmission in each direction on up to 100 meters of fiber optics
- It leverages Intel's silicon photonics technology and a silicon photonics integrated circuit (PIC), which includes on-chip lasers and optical amplifiers, with an electrical IC (EIC).



DGX H100: DATA-NETWORK CONFIGURATION

Full-BW Intra-Server NVLink

- All 8 GPUs can simultaneously saturate 18 NVLinks to other GPUs within server
- Limited only by over-subscription from multiple other GPUs

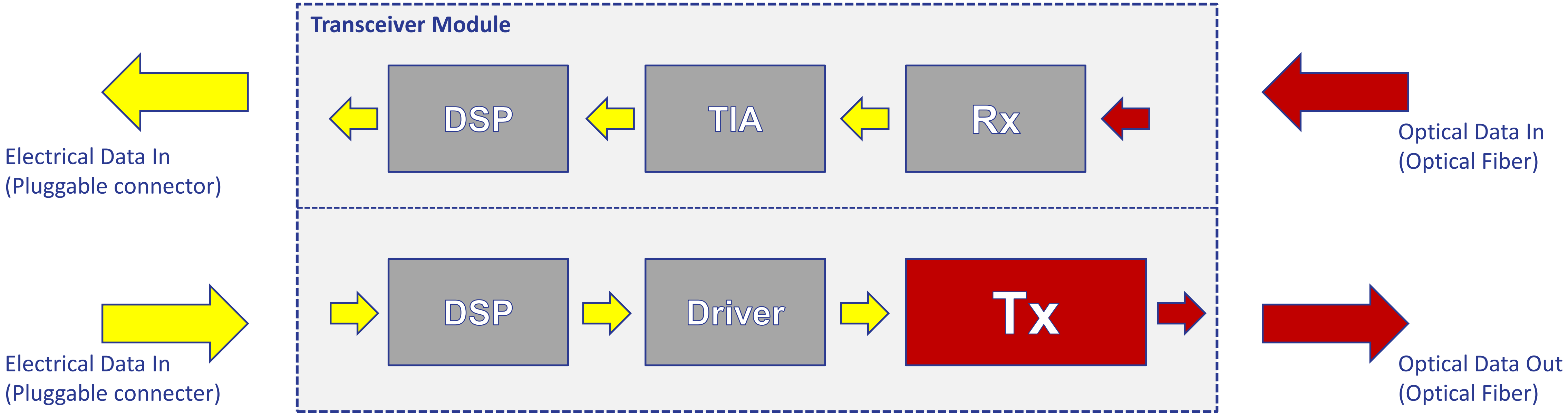
Half-BW NVLink Network

- All 8 GPUs can half-subscribe 18 NVLinks to GPUs in other servers
- 4 GPUs can saturate 18 NVLinks to GPUs in other servers
- Equivalent of full-BW on AllReduce with SHARP
- Reduction in All2All BW is a balance with server complexity and costs

Multi-Rail InfiniBand/Ethernet

- All 8 GPUs can independently RDMA data over its own dedicated 400 Gb/s HCA/NIC
- 800 GBps of aggregate full-duplex to non-NVLink Network devices

Optical Transceiver Overview



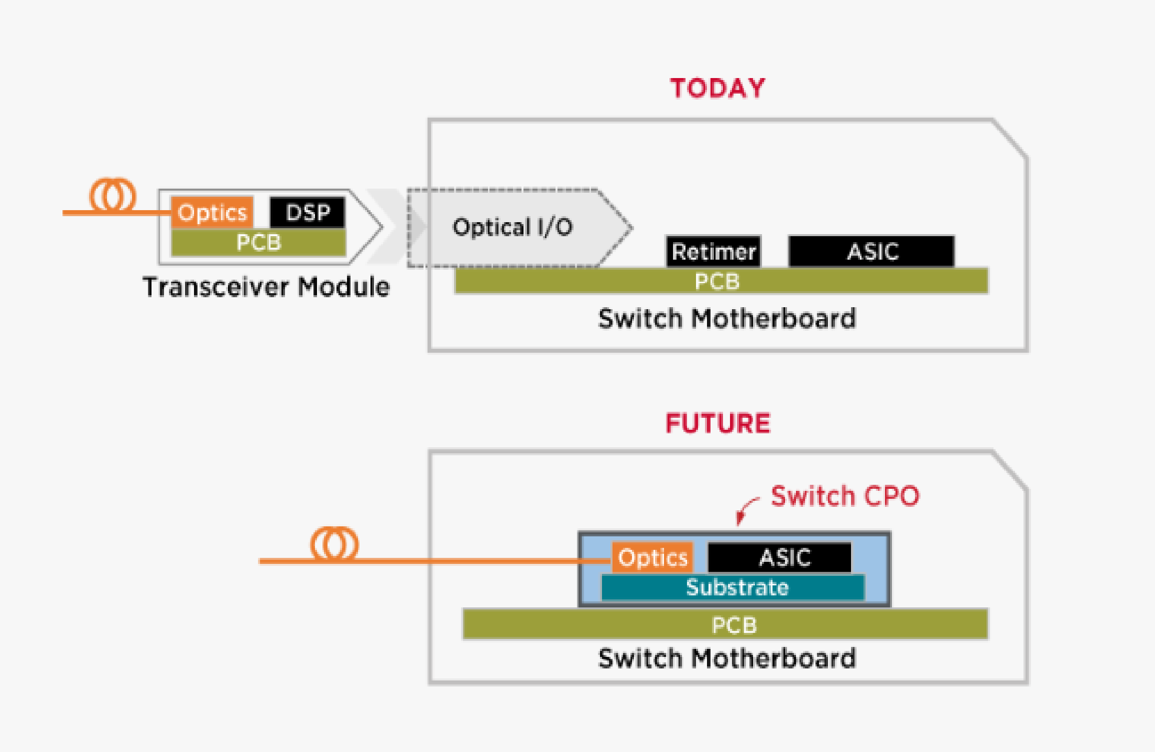
- Transmitter (Tx) function: Data transmit – Take incoming electrical data from the driver and send out as optical data signal
- Receiver (Rx): Data receive – Take incoming optical data from the fiber and convert them as electrical data signal

Co-Packaging Optics

From 100G, 400G, 800G Transceivers to CPO

- Co-Packaged Optics (CPO) is an advanced heterogeneous integration of optics and silicon on a single packaged substrate. CPO brings together a wide range of expertise in fiber optics, digital signal processing (DSP), switch ASICs, and state-of-the-art packaging & test to provide disruptive system value for the data center and cloud infrastructure

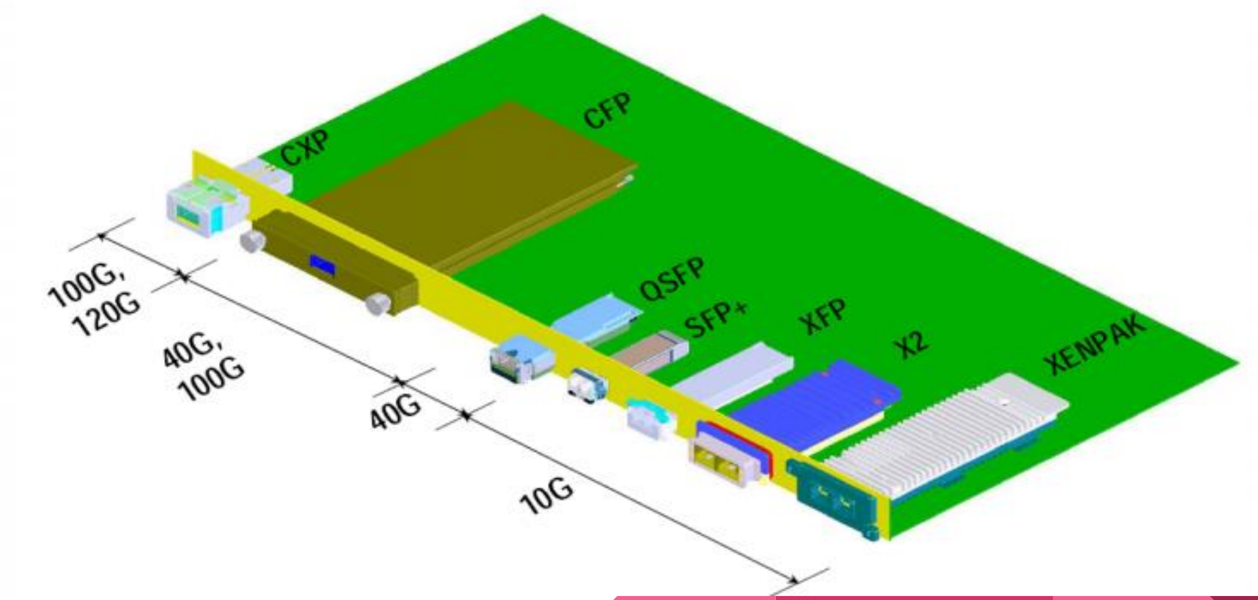
Opportunities for Advanced Packaging



https://www.broadcom.com/info/optics/cpo

Factors for Optical Transceivers

- Speed/Data Rate: 100 Gbps vs 28 Gbaud
- Form Factor: The international standard multi-source agreement (MSA) defines the different transceiver form factors. Popular form factors include SFP, SFP+, XFP, CFP, CFP2, SFP28, QSFP+, QSFP28, QSFP56, QSFP-DD...
- Transmission Distance
  - CR: Copper Reach, 3m
  - VR: Vertical Reach, 50m
  - SR: Short Reach, 100m
  - DR: Data Center Reach, 500m
  - FR: Fast Reach, 2km
  - LR: Long Reach, 10km
  - ER: Extended Reach, 40km
  - ZR: Ze Best Reach, 80km
  - XR: Varible Reach and data rate

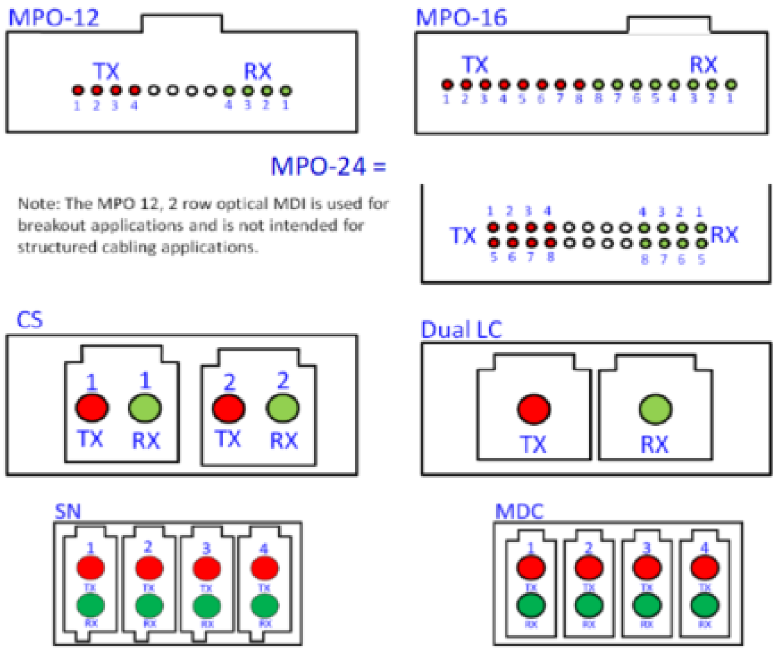


https://www.linkedin.com/pulse/6-tips-choosing-right-transceiver-optics-etern-optoelectronics/

Factors for Optical Transceivers – Cont’d

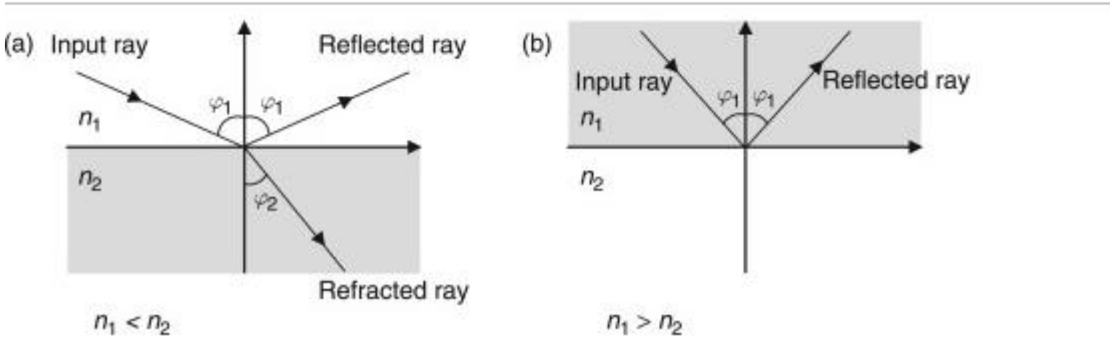
- Fiber Type:
  - Single mode means the fiber enables one type of light mode to be propagated at a time. While multimode means the fiber can propagate multiple modes
- Connector/Optical Interface
  - The most common types of transceiver connectors used in optical transceiver are LC and MPO
- Case Temperature
  - Common applications need only commercial temperature transceivers at 0 to 70C. There is extended (-20 to 85C) and industrial temperature (-40 to 85C) transceivers for outdoor and rugged applications

Fiber Medium	MMF	SMF	SMF
Wavelength/nm	850	1310	1550
Insertion Loss dB/km	3.5	0.35	0.25



Basic Concepts of Optical Fibers

- Optical fibers are basically composed of two coaxial layers: core and cladding
- The principle of light propagation through a fiber is total internal reflection at the simple interface between two different dielectric materials. When light is incident at the boundary, it gets refracted and reflected



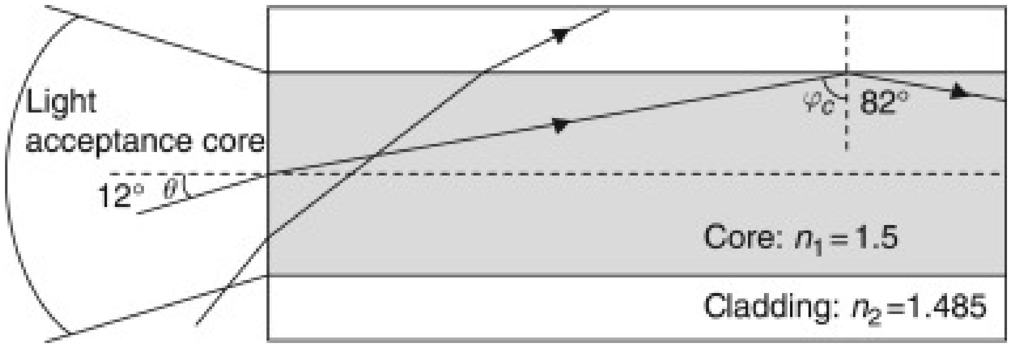
$$\frac{\sin \phi_1}{\sin \phi_2} = \frac{n_2}{n_1}$$

This is called Snell's law. For $n_1 > n_2$, sin φ_2 approaches 1 as φ_1 increases to $\varphi_c = \sin^{-1}(n_2/n_1)$. The angle φ_c is referred to as the critical angle

*https://www.sciencedirect.com/science/article/pii/B9780444533494002090

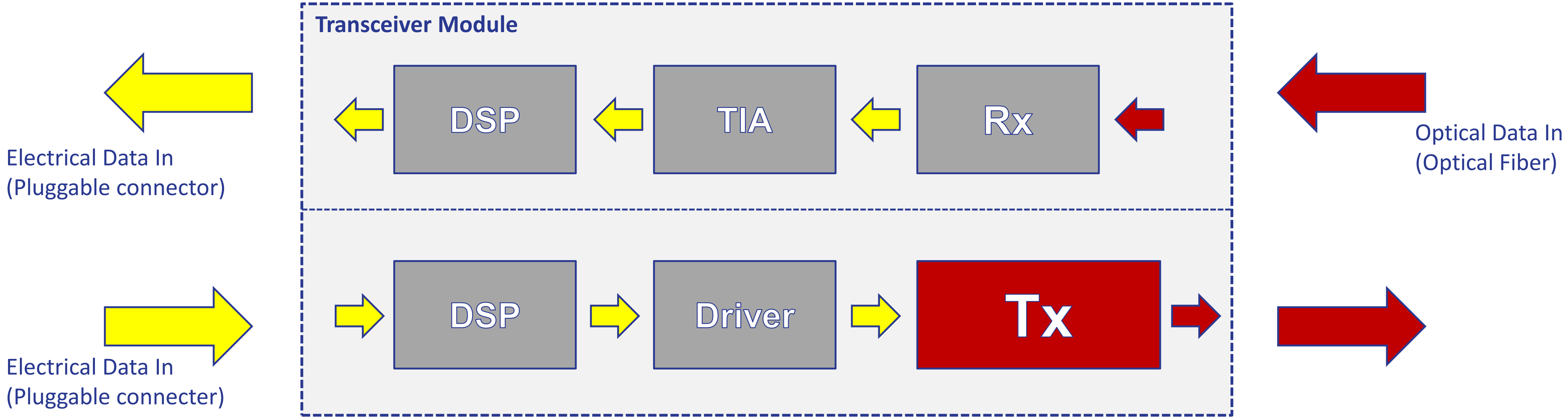
Numerical Aperture (NA)

- The refractive index of the core is slightly higher than that of the cladding.
- If the light ray strikes the core/cladding boundary at a smaller angle than the critical angle, only partial reflection takes place and some of the energy is lost by refraction into the cladding
- The numerical aperture (NA) is generally utilized as an important parameter that describes fiber properties such as coupling loss and bending loss



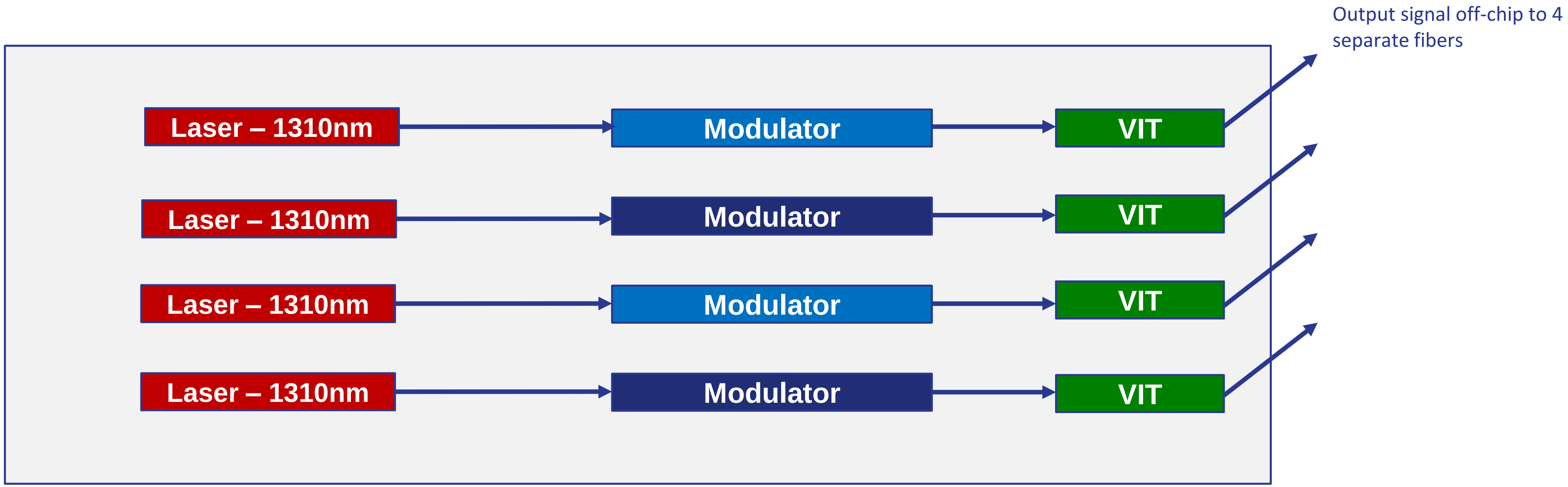
$$\sin\theta = \sqrt{{n_1}^2 - {n_2}^2} = \mathrm{NA}$$

100G PSM4 – Package Architecture



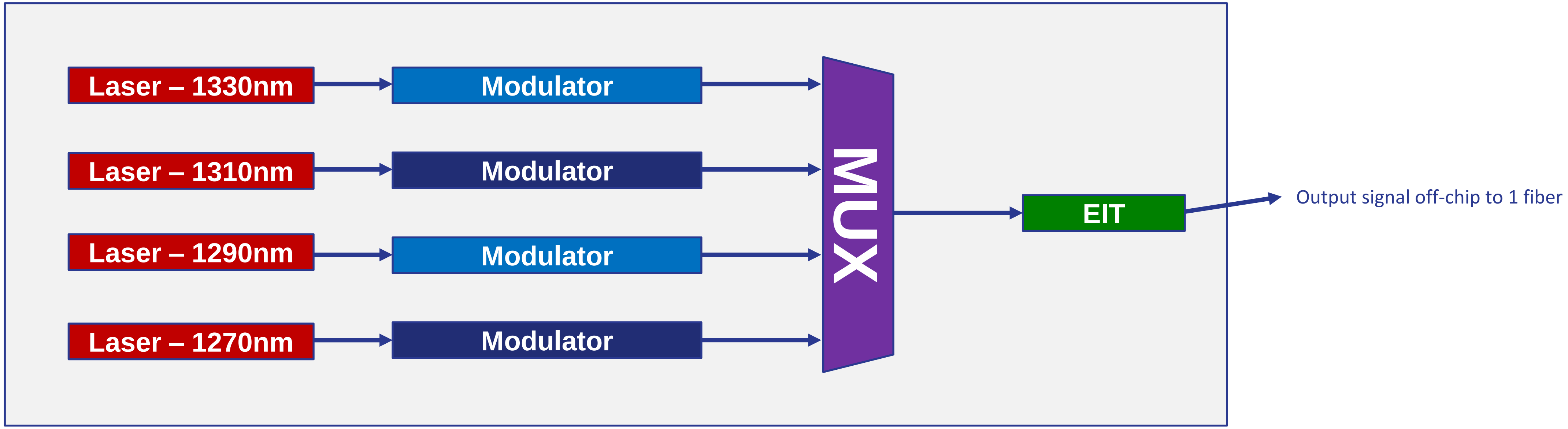
- **1**: High-speed data comes in from the substrate onto the Tx and goes to the driver input (**2**).
- The driver amplifies the input signal to drive the modulator phase shifter (**3**).
- DC traces deliver power to the driver IC (**4**), while laser bias (**5**) and modulator bias control (**6**) signals are also connected through wirebond pads to the substrate.
- The modulator encodes the data onto the optical signal, which is coupled off-chip through the VIT (**7**).

100G PSM4 – Component Architecture



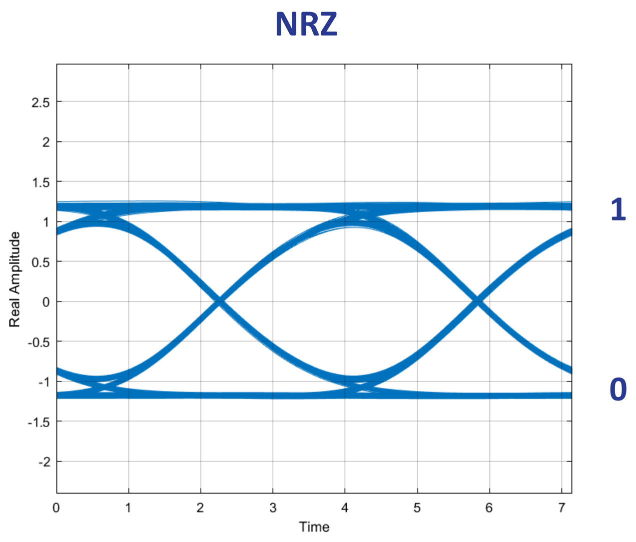
- PSM4 (Parallel Single Mode 4)
- 4 parallel channels (separate fibers) x 25Gbps = 100G
  - **Laser**: Generate light
  - **Modulator**: Modulate, encode data to the optical signal
  - **VIT**: Couple light off-chip and into fiber

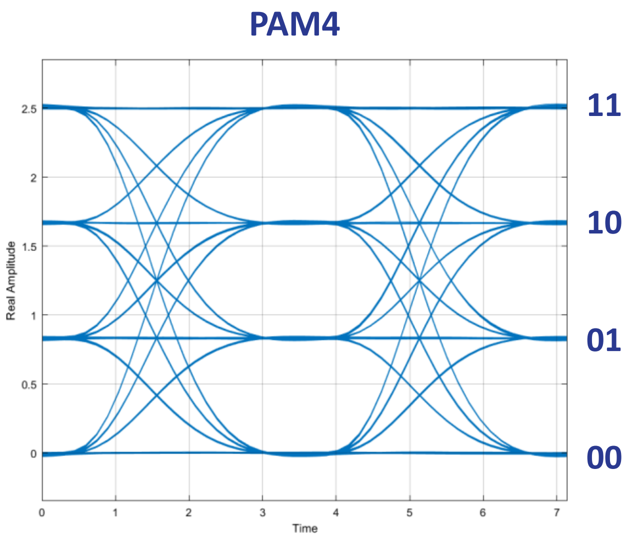
100G CWDM4



- 4 separate wavelengths x 25Gbps x 1 fiber = 100Gbps
- The key difference for CWDM4 compared to PSM4 is that the channels are multiplexed (MUX) by wavelength into a single fiber, rather than separate in 4 parallel fibers
- Product variants: 500m, 2km, 10km for inside-datacenter, 5G wireless for metro/cellular applications

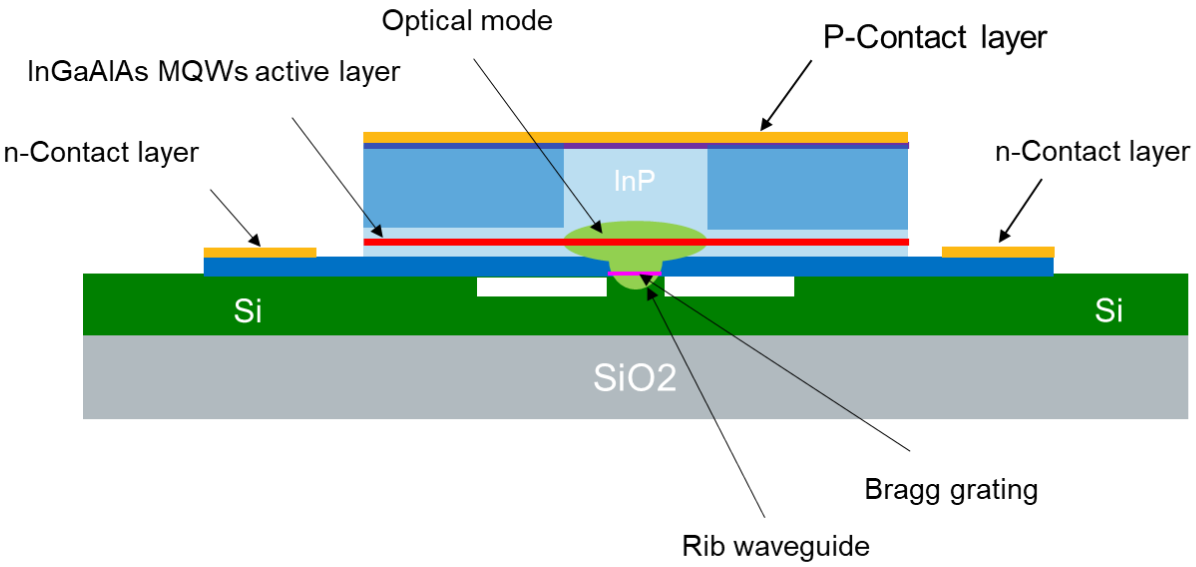
400G DR4





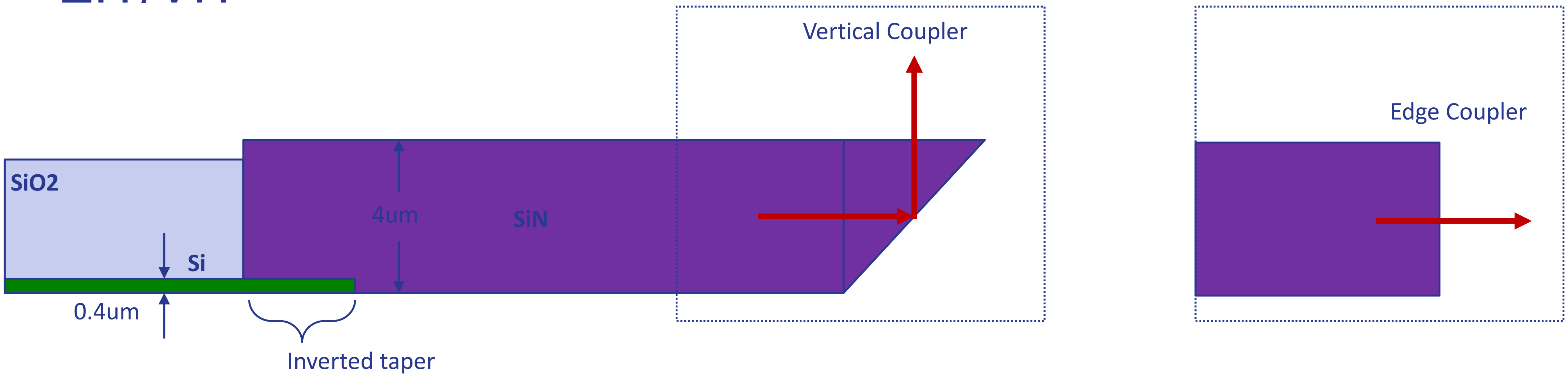
- Based on **PAM4** modulator (pulsed amplitude modulation, 4 levels)
  - 4 levels of data → 2x the bits per symbol
- DR**4**: Next generation of PSM4
  - 50GBaud/s PAM4 per channel →100Gbps per channel x **4** channels parallel = 400G

Laser



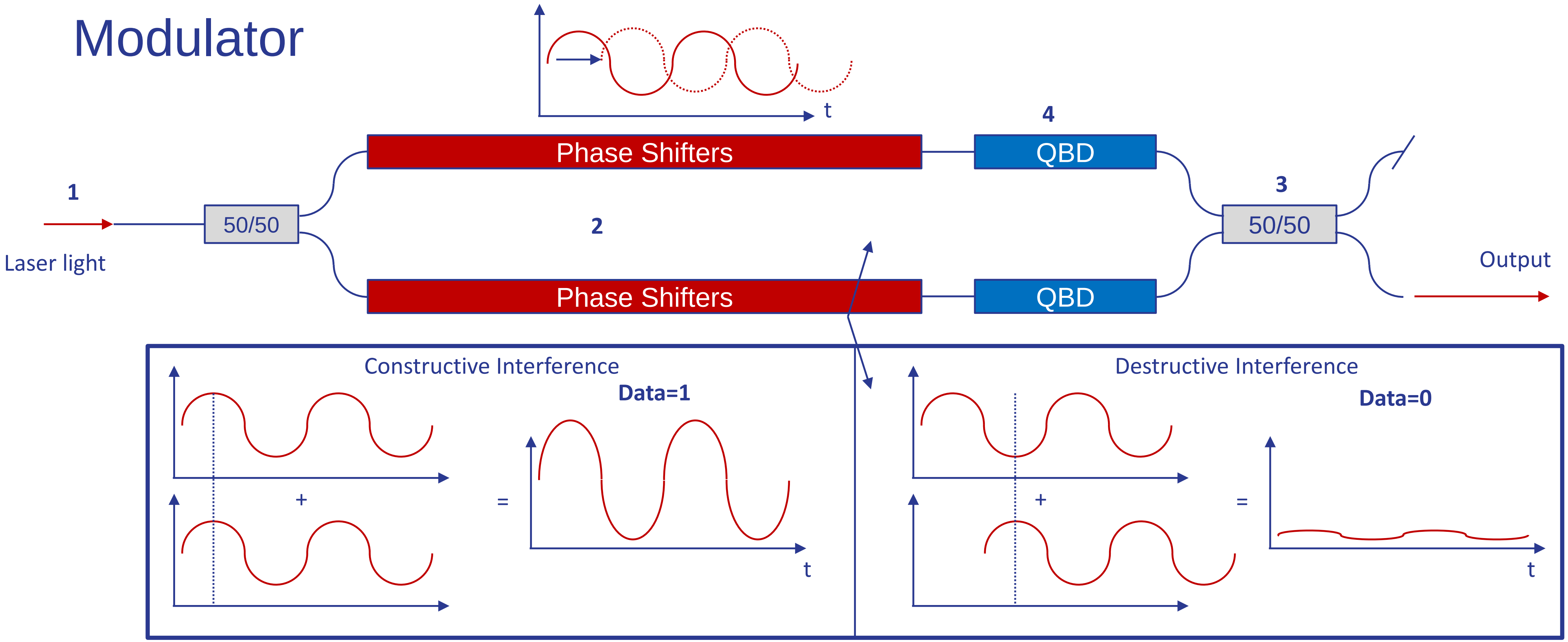
- Silicon itself cannot be used to make electrically-pumped lasers.
- **III-V/Silicon Hybrid Laser**: Wafer bond InP to silicon and fabricate integrated on-chip lasers. InP provides light emission and amplification for lasing. Silicon provides light confinement.

EIT/VIT



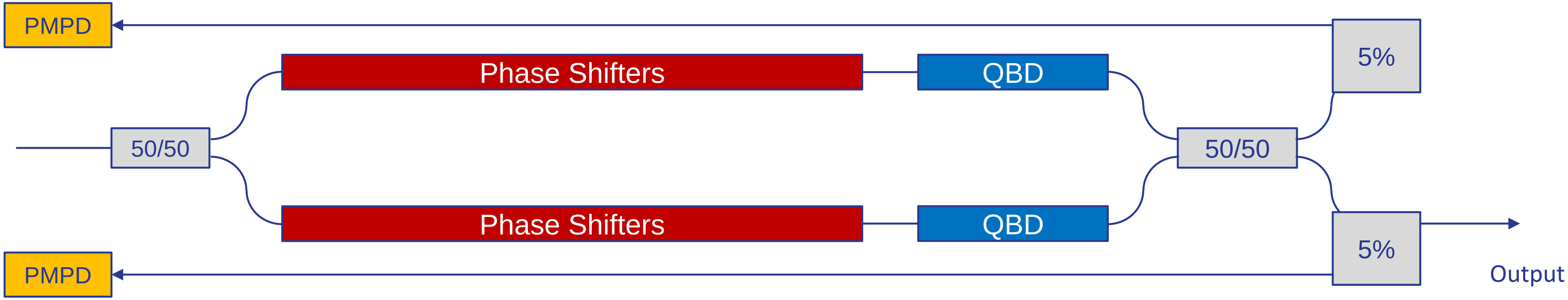
- **EIT** = Edge Inverted Taper, **VIT** = Vertical Inverted Taper
- The IT transitions the light from a small sub-micron sized optical mode to a larger mode in a silicon nitride waveguide.
- For a **VIT**, the light is reflected vertical out of the surface of the chip w/ a 45° mirror. For the **EIT**, the light is emitted out of the edge facet of the chip.

Modulator



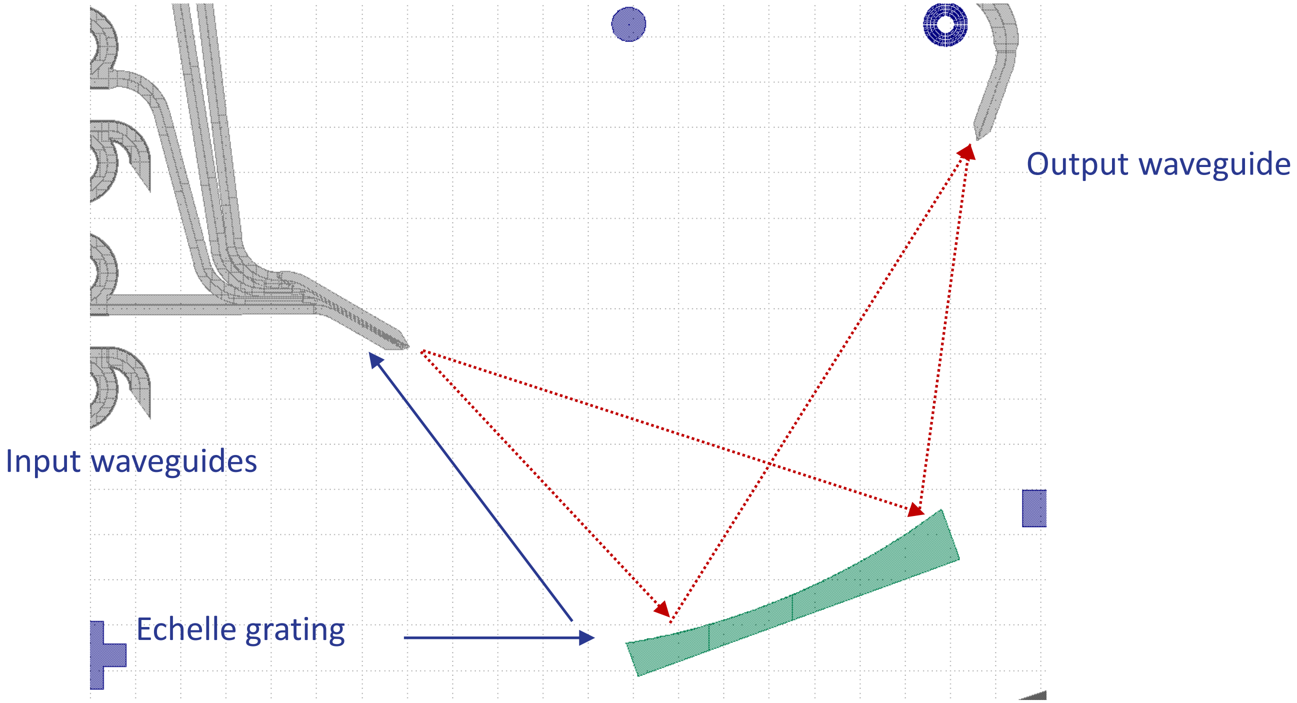
- **Silicon MZI (Mach-Zehnder Interferometer) Modulator**
- Light is split into 2 arms (**1**), phase shifted (**2**) and recombined at the output 2x2 coupler (**3**). The constructive and destructive interference of the phase-shifted light creates the 1's and 0's to modulate and create encode the digital data onto the optical signal.
- The quadrature bias diode (QBD) (**4**) sets the DC bias point for the modulator.

Modulator – Feedback Control



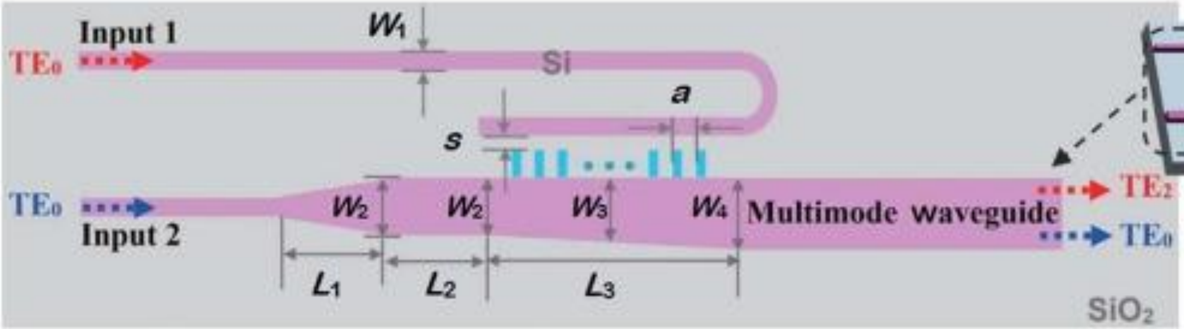
- 5% taps split off from each arm of the modulator and fed back to power monitor photodetectors (**PMPD**), which monitors the optical power through the modulator to set the DC bias point as well as the laser bias current.

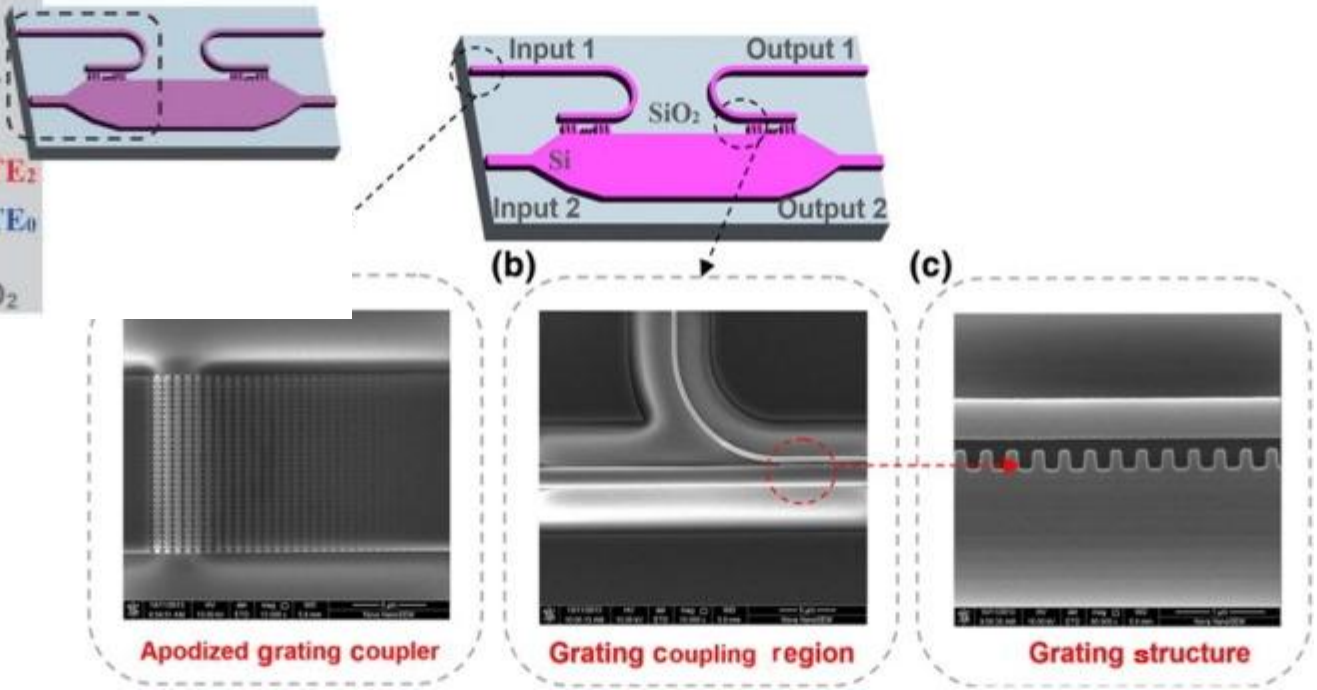
MUX



- Light goes from the input waveguide and is reflected by the echelle grating, which acts as a mirror and reflects the light to the output waveguide.
- Since the refractive index of silicon is wavelength dependent, the angle of reflection off the echelle grating is wavelength dependent.
- The input waveguides are positioned in the optimal location to match the wavelength of each channel.

Mode (De)Multiplexing Waveguide





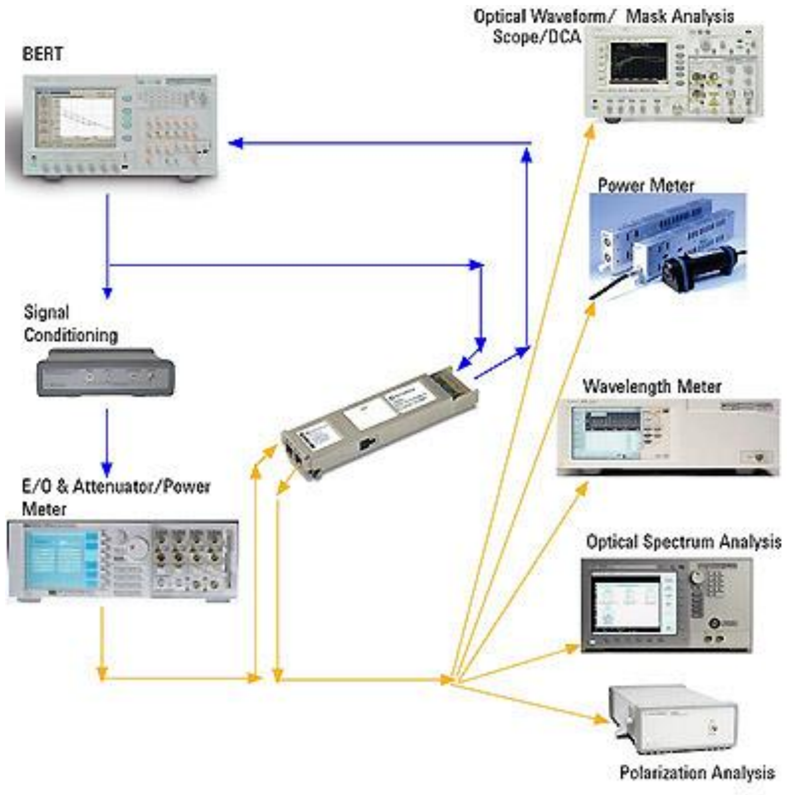
Through a grating-assisted directional coupler, the upper narrow waveguide is coupled with a bus waveguide. In the bus waveguide, TE_2 mode is excited owing to the mode conversion. TE_2 carries the information of TE_0 mode launched from Input 1. After the region of grating-assisted directional coupler, TE_0 mode and TE_2 mode propagate simultaneously along the multimode waveguide. Hence, the mode multiplexer is achieved by adopting the grating-assisted directional coupler. Similarly, the demultiplexer can be also achieved by another grating-assisted directional coupler with the same parameters.

https://www.sciencedirect.com/science/article/pii/S209592731830327X

Testing overview

Front-end optical wafer testing

Back-end transceiver module testing

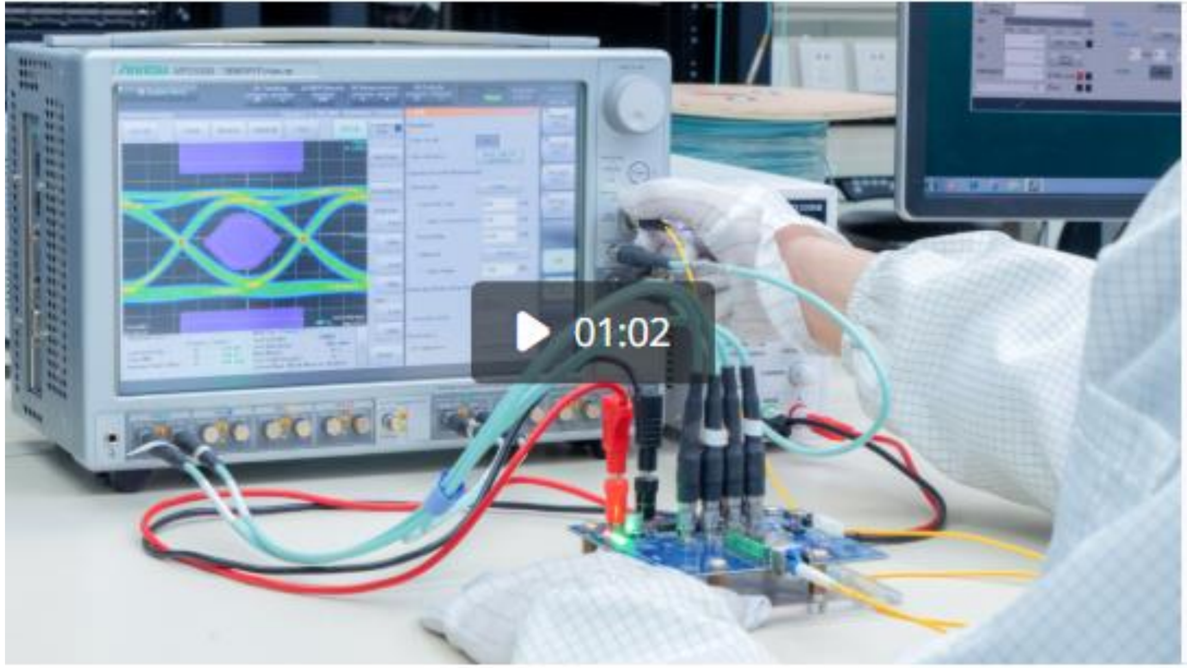


Power and Wavelength Testing

Test the signal delivering strength and wavelength, to ensure the signal decoding capacity of the receiver, and the wavelength remains consistent from the transmitter to the receiver.

Traffic Testing

Test the bit error rate and packet loss rate, to make them meet the corresponding standards and ensure the performance of transceivers.



Optical Performance Testing

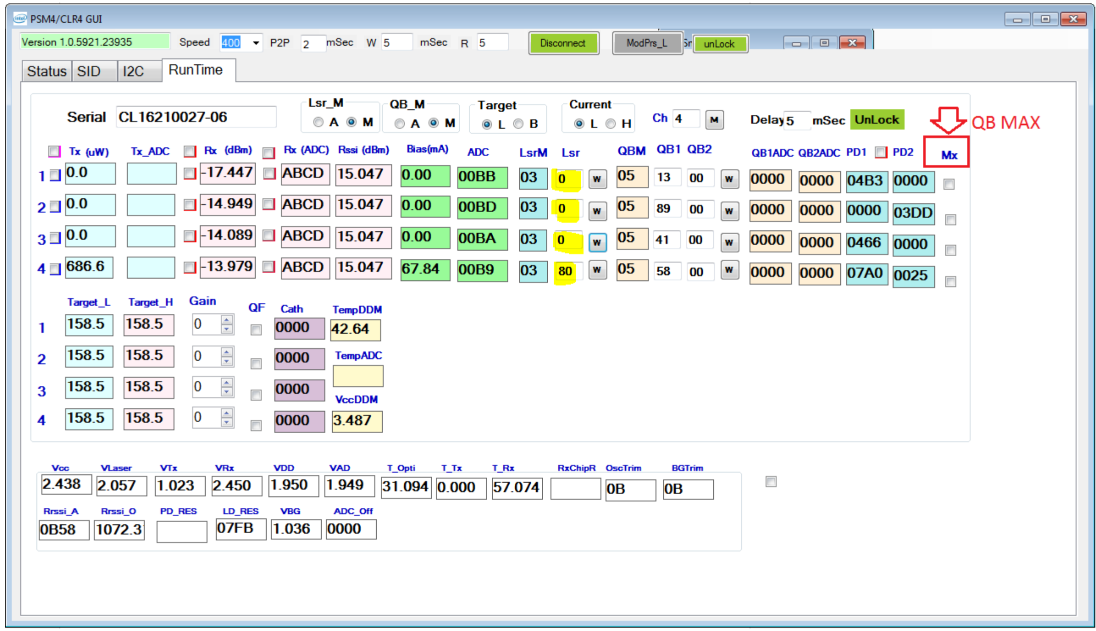
Test the transceivers' eye diagram situation, receiving sensitivity, extinction ratio, wavelength, light-emitting, light-receiving, current and voltage, to ensure the signal quality, stability and reliability of the transmission.

End Face Testing

Check the end face of the transceivers and keep them clean for more stable data transmission, better performance, and durability.

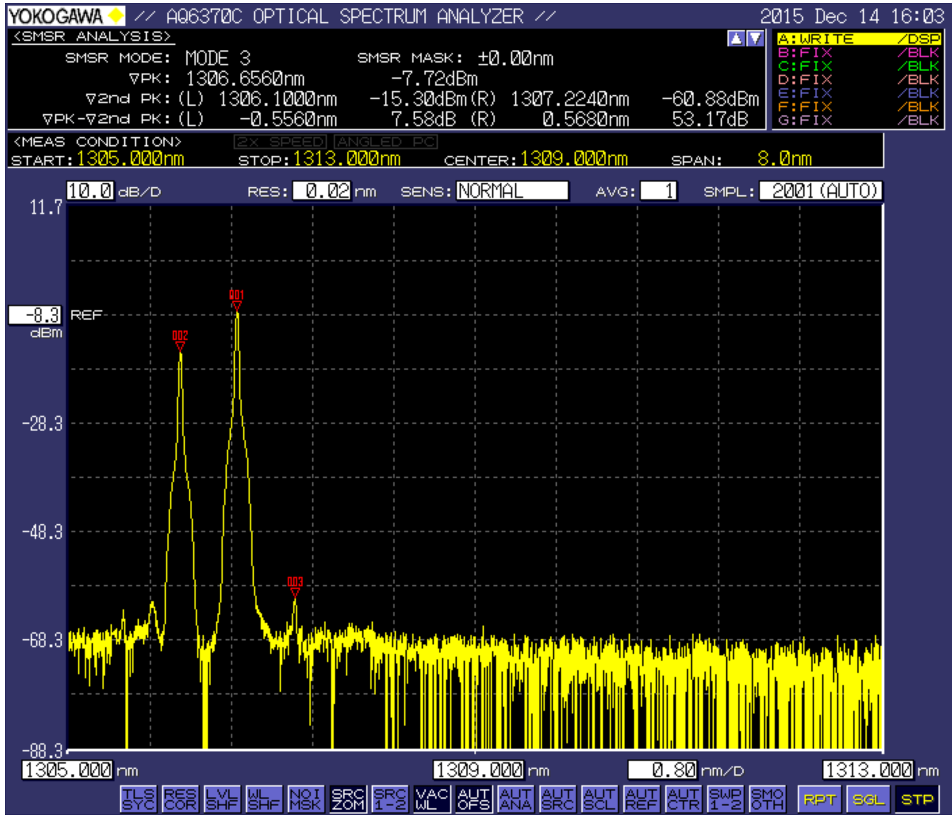
DDM Function

- **DDM**: “Digital Diagnostics Monitoring”. It can monitor the parameters of fiber modules like optical output (TX) power, optical input (RX) power, temperature, laser bias current, and transceiver supply voltage, in real time
  - If there is any problems with your optical transceivers and can not work well, the first thing you get to see all these parameters of your transceivers at that time with DDM.
  - All these “analog” parameters are digitized
  - Controlled by Microcontroller & FW automatically



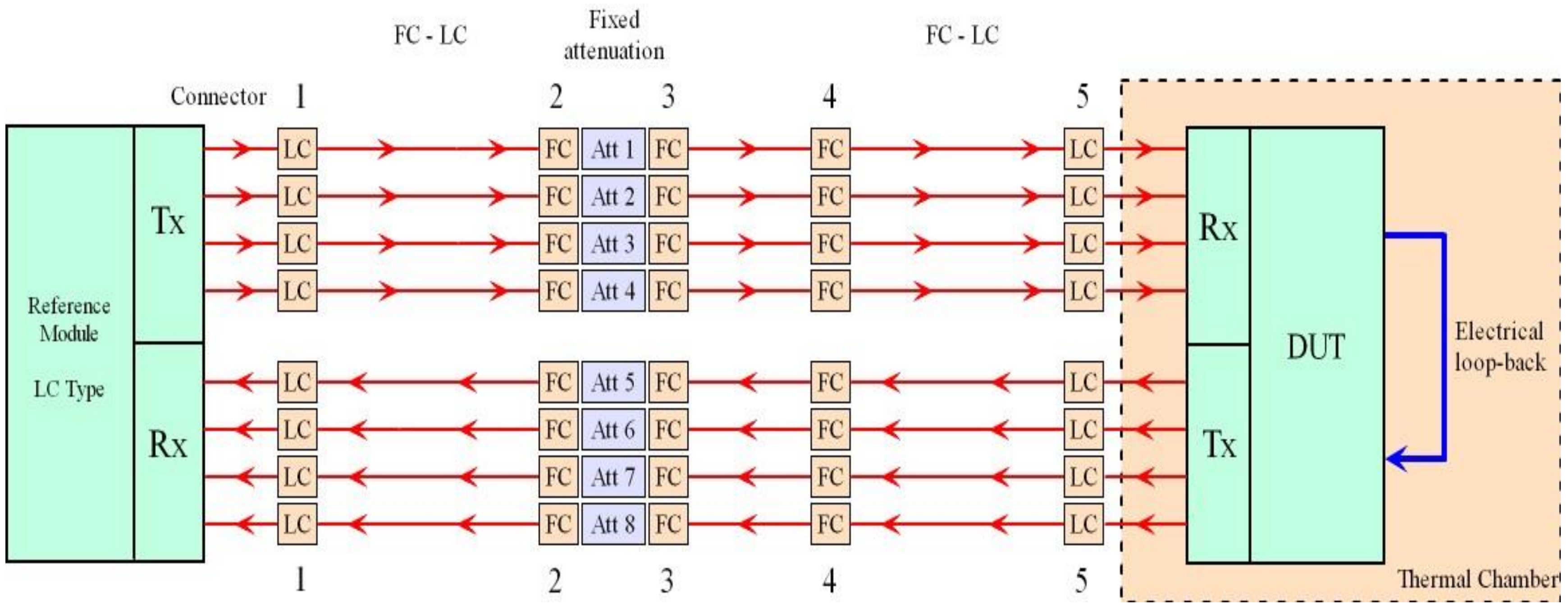
Side Mode Suppression Ratio (SMSR)

- A laser's side-mode suppression ratio (SMSR) describes the amplitude difference between the main mode and the largest side mode in decibels. A typical value is greater than 30 dB, indicating most of the power exists in the main mode
- Mode hopping refers to a phenomenon where a laser system abruptly switches from one longitudinal mode to another. In a laser, different longitudinal modes represent distinct allowed frequencies or wavelengths of the laser cavity.



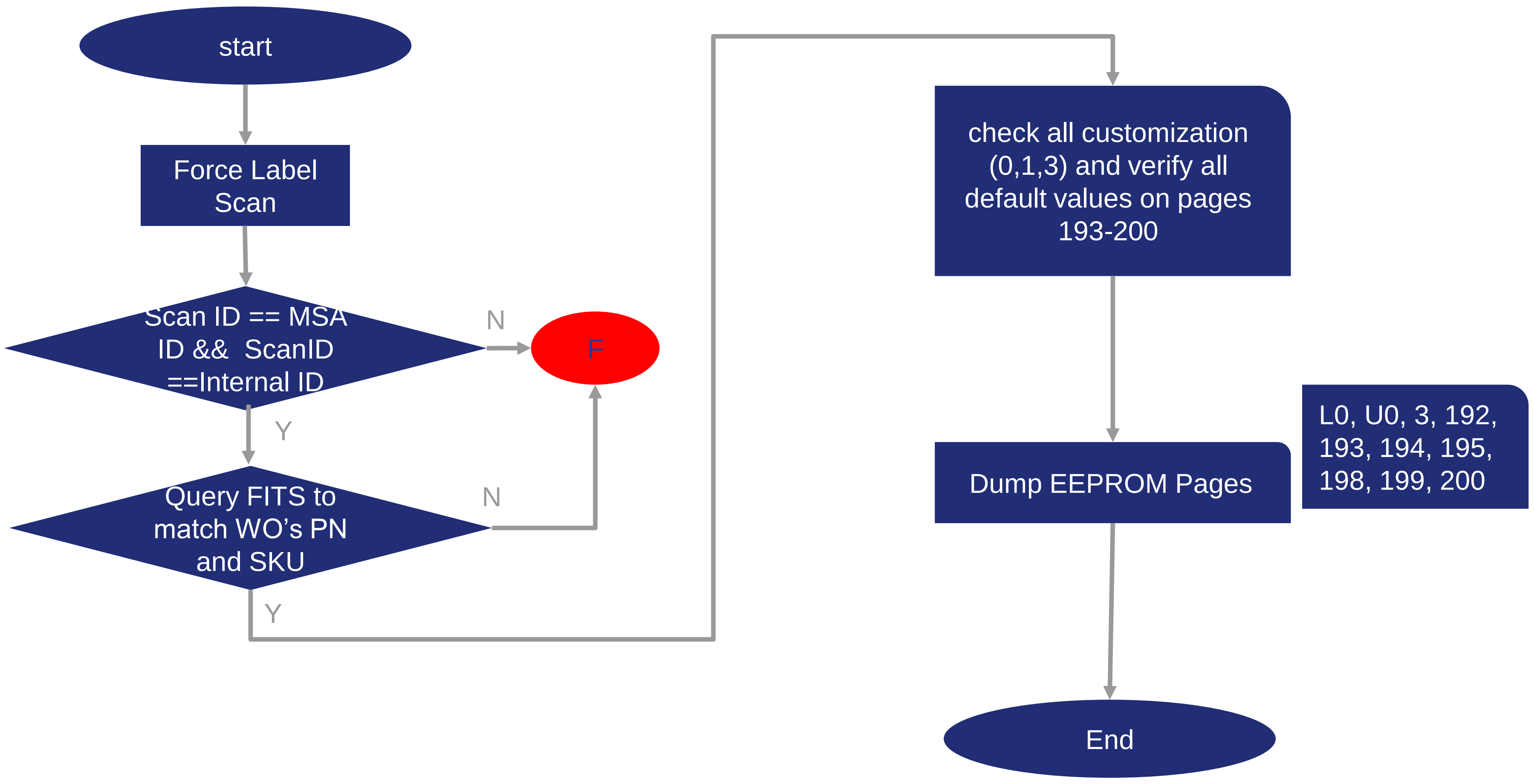
Link Test & Outgoing Test

- Link Test: It usually done with connected Reference golden modules to modules under test 1x1 with electrical loop-back in the Thermal Chamber with temperature ramping up & down. The P/F criteria includes BER < 1e-9 sampled every 2 mins and no LOL (loss of lock or polarity flip)



Outgoing Test

- Ensure the Manufacturing ID matching with ID programed inside FW and Label
- Verify the FW data integrity
- Verify the FW calibration pages
- Dump EEPROM pages as snapshot



AI/ML Application – Data Feed Forward and Backward

- It is very expensive and time consuming to raise temperature up & down during testing of a DUT -> test entire wafer or lot at the same temperature and then move the DUTs to another temperature vs.. multiple insertions
- The formulae might not be linear and required volume characterization data to model them -> machine learning
- Data Feed Backward: e.g. Monitoring test patterns v.s. defect detection -> if there is no any defects detected on said 1 million DUTs-> the pattern can be removed for TTR
- Think about how to utilize data has collected across all the locations during Assembly and Test

AI/ML Application

- Many AI/ML applications were used for test time reduction -> save the test cost
- Could data collected at Cold (Hot) temperature used to predict the measurement results at Hot (Cold) temperature, so test at Hot (Cold) temperature can be skipped?
- Or, only need to minimize the number of DUTs of False Positive and False Negative
  - A false positive is when something is true when it is actually false (also called a type I error). A false positive is a “false alarm.” -> yield loss
  - A false negative is saying something is false when it is actually true (also called a type II error). A false negative means something that is there was not detected; something was missed. -> test escape (Bad)
- Or, can re-measure these DUTs only -> other Challenges
- Machine learning model build* (training data set vs. test data set)
- Domain knowledge base (Measure a few parameters to calculate complex parameters with characterization data set)

“Machine learning application for silicon photonics transceiver testing,” Woosung Kim, Yeoh Hoe Seng, Yi-Shing Chang, Suohai Mei; SPIE 2018, San Diego, California, United States

Data Disaggregation

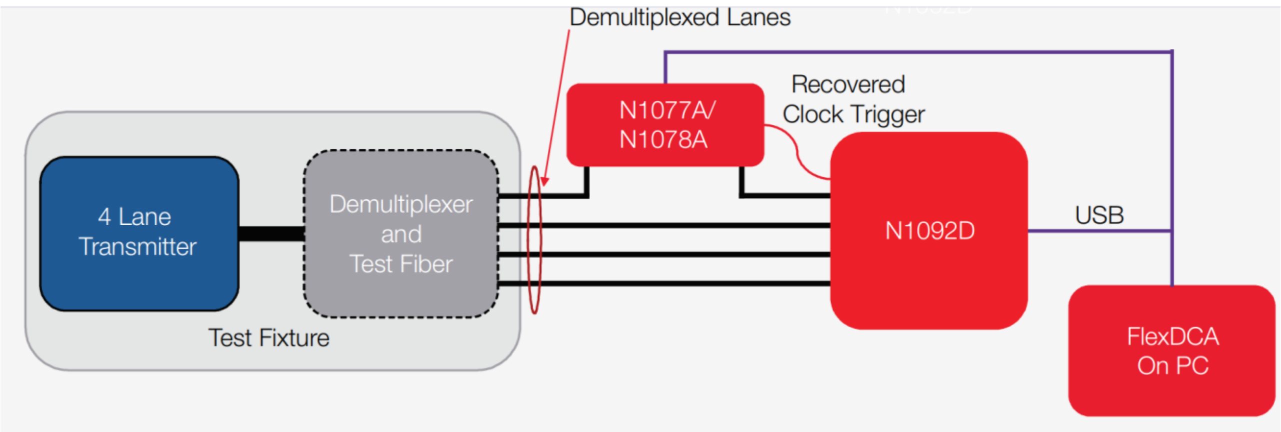
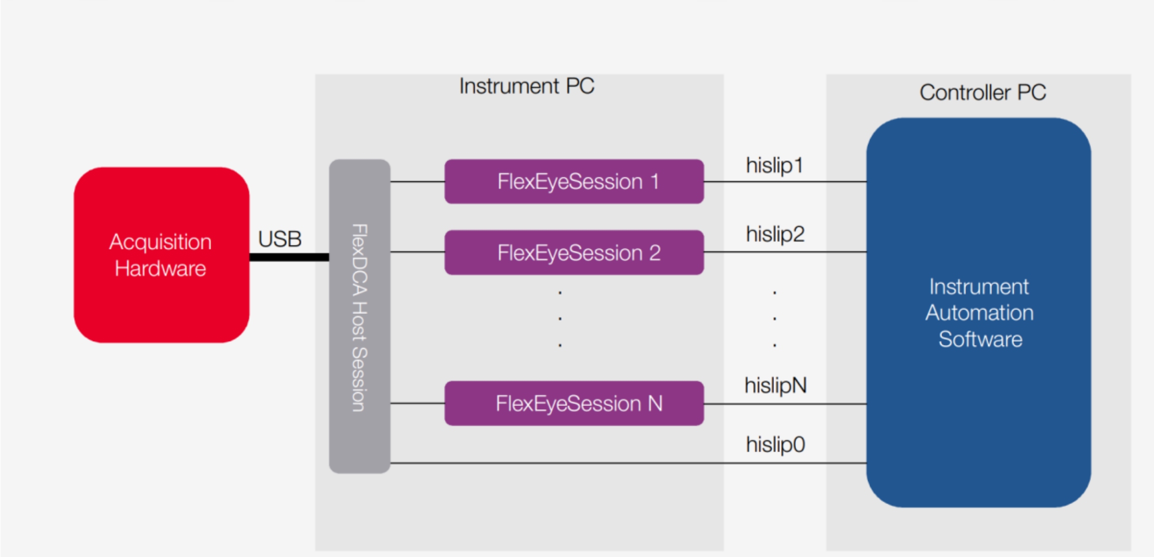
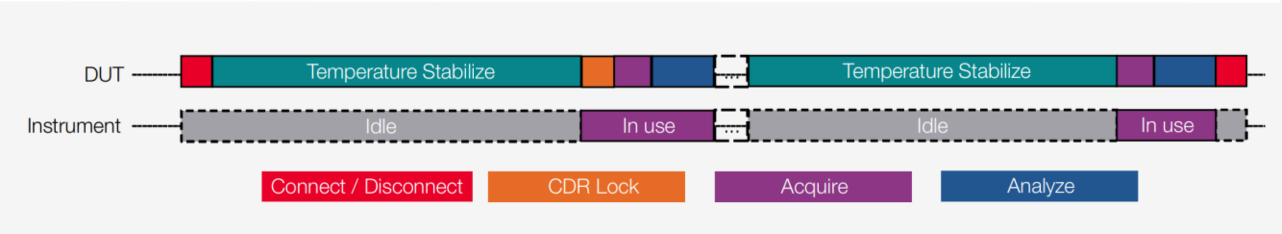
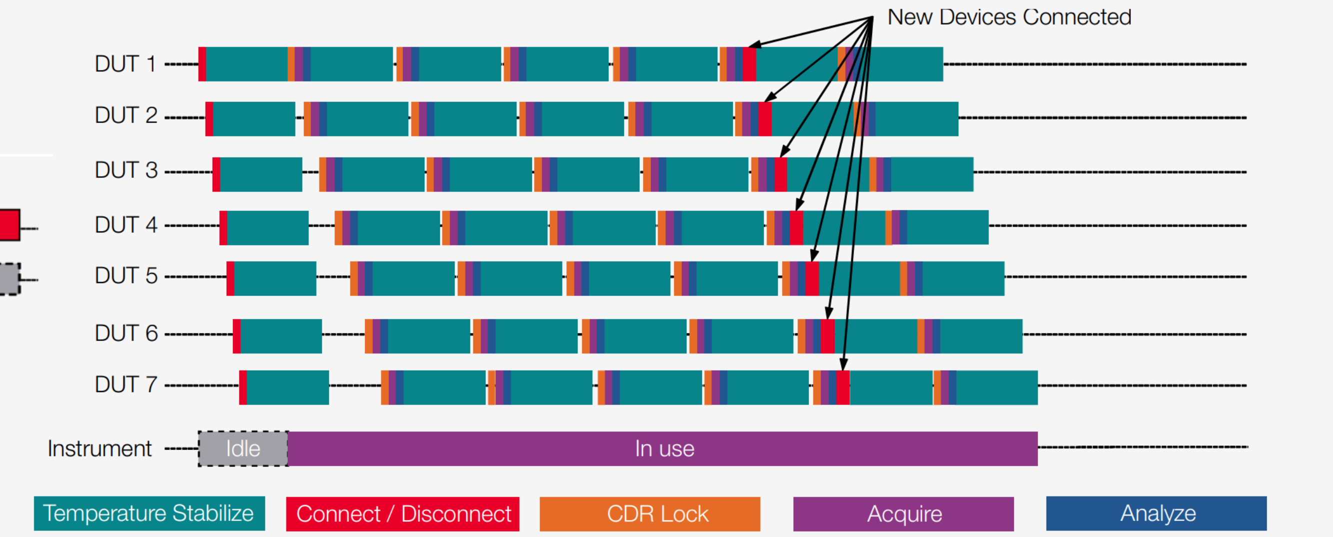


Figure 2. Typical 4-Lane Test Setup with CDR Module







Keysight white paper TDECQ Part II: https://www.keysight.com

Summary

AI Systems in Datacenter Drive Optics Growth

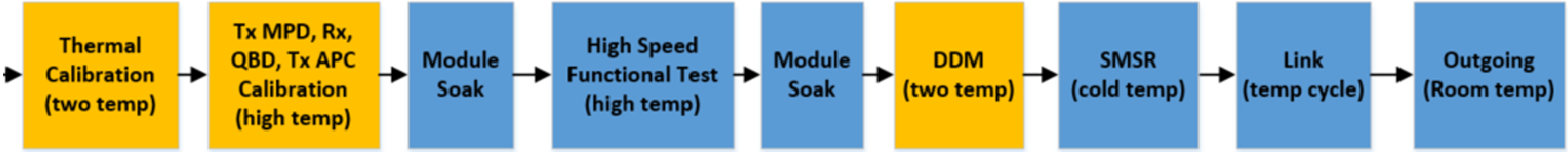
Silicon Photonics Product Overview: from 100G to 800G, to CPO

Optical Transceiver Module Test Overview

AI/ML Application

Backup

Optical Transceiver Testing Process



Thermal Calibration: Laser is very temperature sensitive. Make sure the laser temperature reading is accurate

DDM (Digital Diagnostics Monitoring): monitors the parameters of optical modules like optical output power (Tx), optical input power (Rx) , temperature, laser bias current, and transceiver supply voltage, in real time with digitized values

Tx MPD (Monitor PD), Rx, QBD, Tx APC Calibration at Hot (worst case corner)

High Speed Functional Test (Tx Eye, Rx Bit Error Rate, Rx Sensitivity)