OCP SPECIAL FOCUS:
PHOTONICS

System-Level Evaluation of XPO for High-Density Optical Interconnects in AI Infrastructure

Daniel Mohaghegh *Hardware Engineer, Microsoft*

Sunil Priyadarshi *Senior Director, Arista*

OCP PLATINUM Member

Agenda

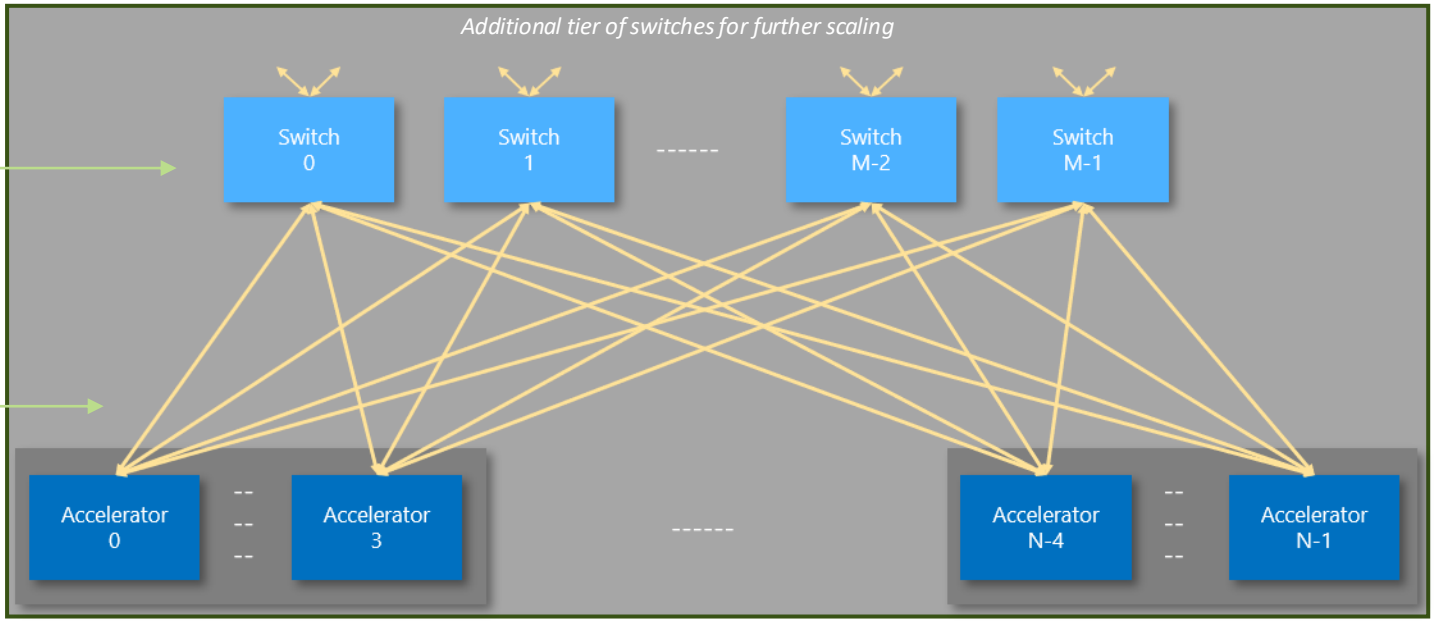
- XPO High-Density Pluggable Form Factor, as applied to AI Scale-Up
- Detailed XPO platform Test Data
- Q&A

XPO High-Density Pluggable Form Factor, as applied to AI Scale-Up

AI Scale-up System Topologies

High Radix switches in a Single Tier enable uniform communication domain

Reach, Bandwidth and Connectivity enabled by optical interconnects in Accelerator trays



Example Topology

Richly connected, Multi-Rack and Multi Row Scale-up Domains

Several topology variations are possible depending on target AI workload performance and cost

Scale-up Interconnect Technology Evaluation: Towards at-scale deployment

Ph 1: Ingredient Readiness
Prioritize technologies
Integration complexity
Link performance and power

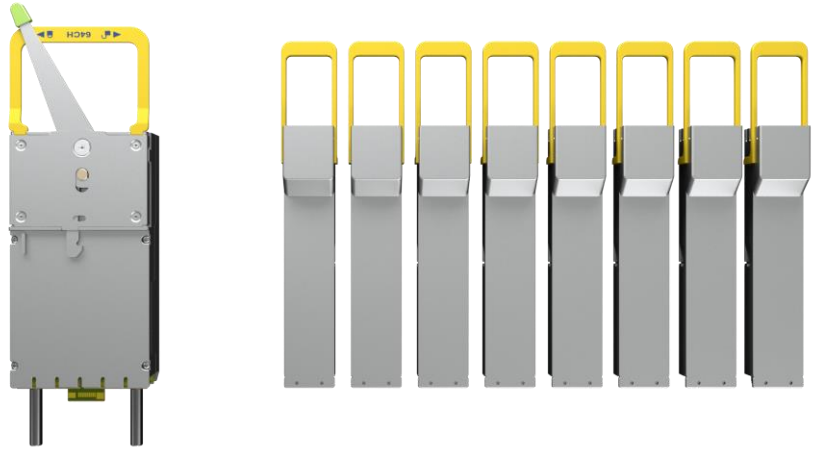
Ph 2: System Design Readiness
End-to-end link performance
Optical device reliability
Accelerator/Switch/Rack design
Fiber management
Ecosystem / Supply chain

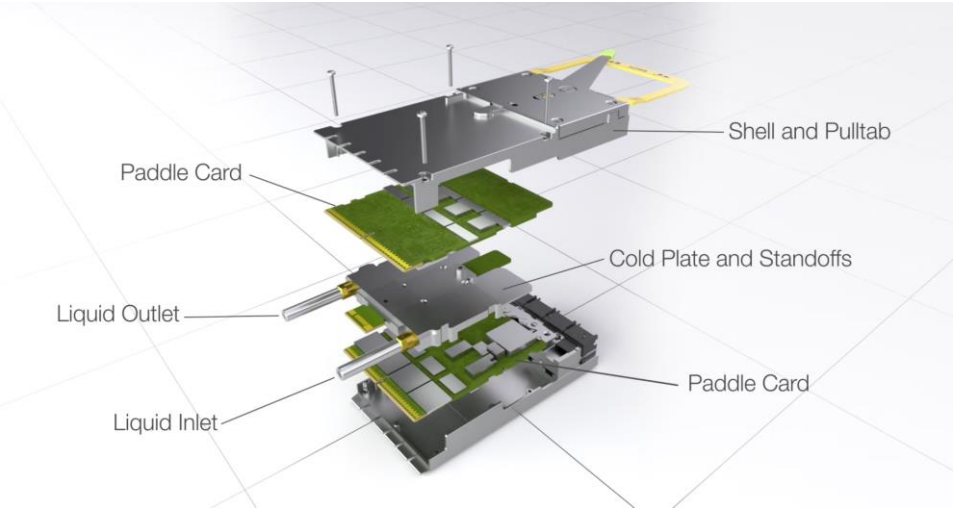
Ph 3: Deployment Readiness
DC installation & operations
RAS
Telemetry
Firmware and Manageability
Control plane software
Data plane software

Increasing scope, down-selection of technologies, 2-3 year span per generation

Intro to XPO

1. Highest Density Pluggable Optics form factor in the Industry
2. 204.8T per 1OU with 200G/lane, 409.6T per 1OU with 400G/lane
3. Supports all existing and emerging optics and interconnect standards
4. Liquid cooling up to 500W per module (for 25.6T coherent-lite)
5. Much lower failure rates with fewer components and lower temperatures
6. Vibrant XPO MSA ecosystem supported by all optics module vendors





XPO
OPTICS INNOVATION

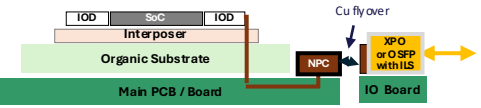
XPO MSA Core Optics Module Members

Accelink Adtran Amphenol APERION
ARISTA ciena COHERENT eoptolink
FIT Genuine optics LIGENT LIGHTMATTER
Linktel LUMENTUM LYNTERA MARVELL
molex O-Net Technologies SOURCE PHOTONICS TERAHOP

Industry Directions on Accelerator-Side Optical Module Form Factors

Pluggable Optics

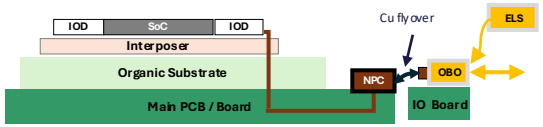
XPO or OSFP



Power	LPO, TRO, FRO
Design Scope	Tray
Serviceability	Module-level
Blast Radius	Module
Standards	Yes

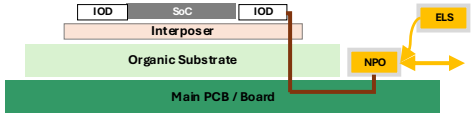
On-Board Optics (OBO)

Separate (Off-) Board: OBO



LPO, TRO, FRO
Tray
Tray-level
Module
No

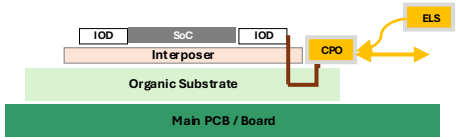
SoC Board: Near-Packaged Optics (NPO)



Linear optics
SoC Board
Tray-level
Module(with CPX)
Open CPX

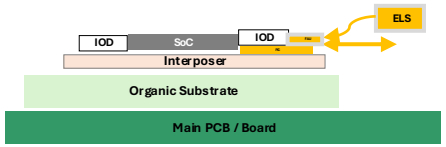
Co-Packaged Optics (CPO)

CPO Multi-chip Module (MCM)



Low
SoC Package
Tray-level
Module & Accelerator
No

CPO Interposer



Very low
SoC Interposer
Tray-level
Module & Accelerator
No

Inference workload bandwidth per SoC is not requiring deeper integration for 1-2 generations

XPO provides Serviceability, Low Blast Radius and Flexibility as a broad industry standard

Accelerator board constraints may encourage exploration of XPO for scale-up optics

Serviceability: The granularity at which a failed optical module can be serviced.
Blast Radius: Cost of replacement (module: only module has to be replaced; accelerator: module and associated accelerator has to be replaced)
Above blast radius definition is in terms of cost of hardware replacement. It is not equivalent AIR or Blast Radius on workload(s)

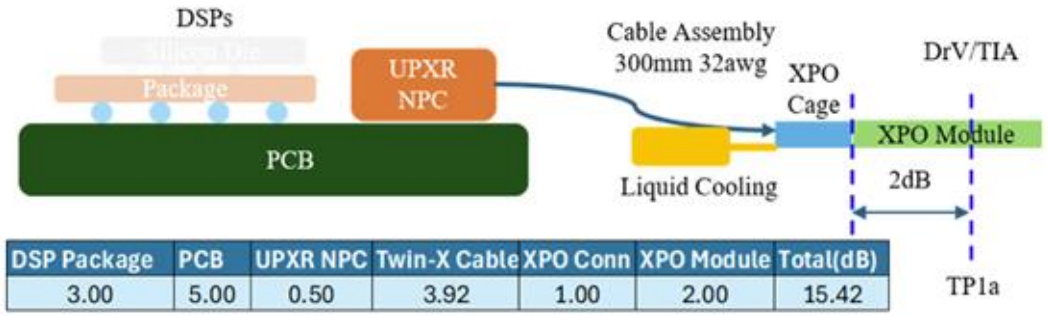
Detailed XPO platform Test Data

XPO Module Test Setup

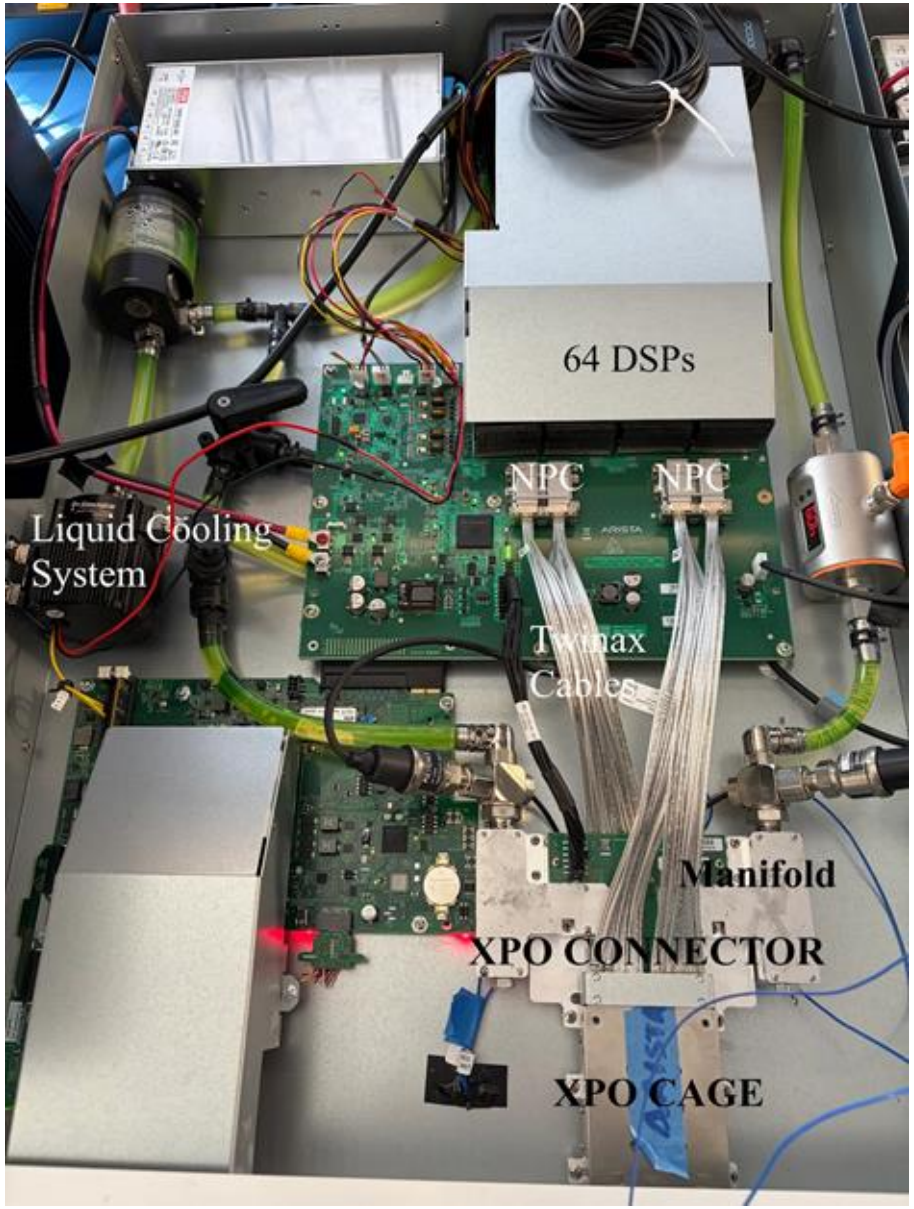
An emulation box with 212 Gbps SerDes was designed by Arista to exercise the 64 channels on the XPO modules.

This emulation platform featured the below characteristics.

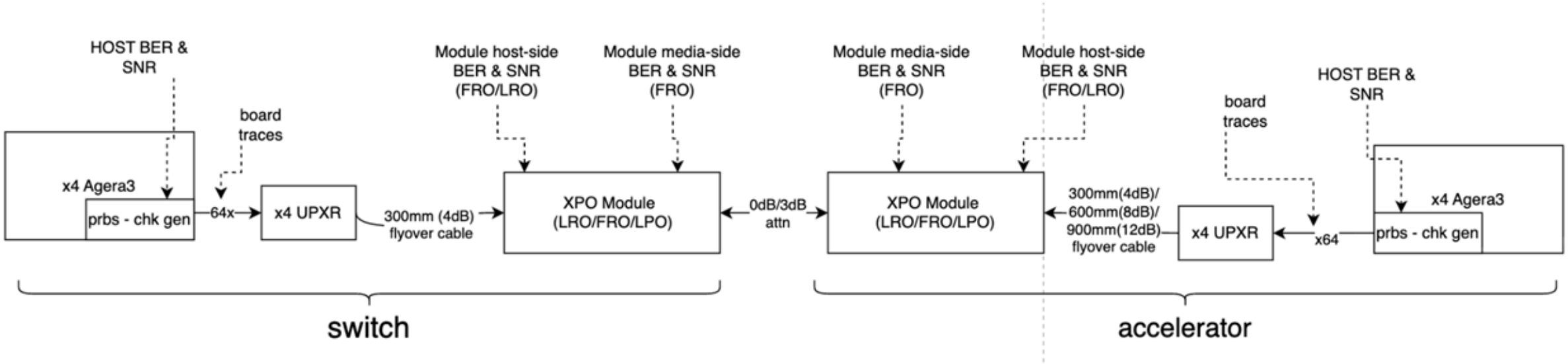
- **x4 Agera3 Broadcom Retimers**
  - x16 lanes 212 Gbps each
- **NPC Copper Flyover Cables**
  - 300/600/900 mm long
  - 32 AWG Twinax: Amphenol XPO to Ultrapass-XR
- **Integrated Liquid Cooling**



DSP Package	PCB	UPXR NPC	Twin-X Cable	XPO Conn	XPO Module	Total(dB)
3.00	5.00	0.50	3.92	1.00	2.00	15.42



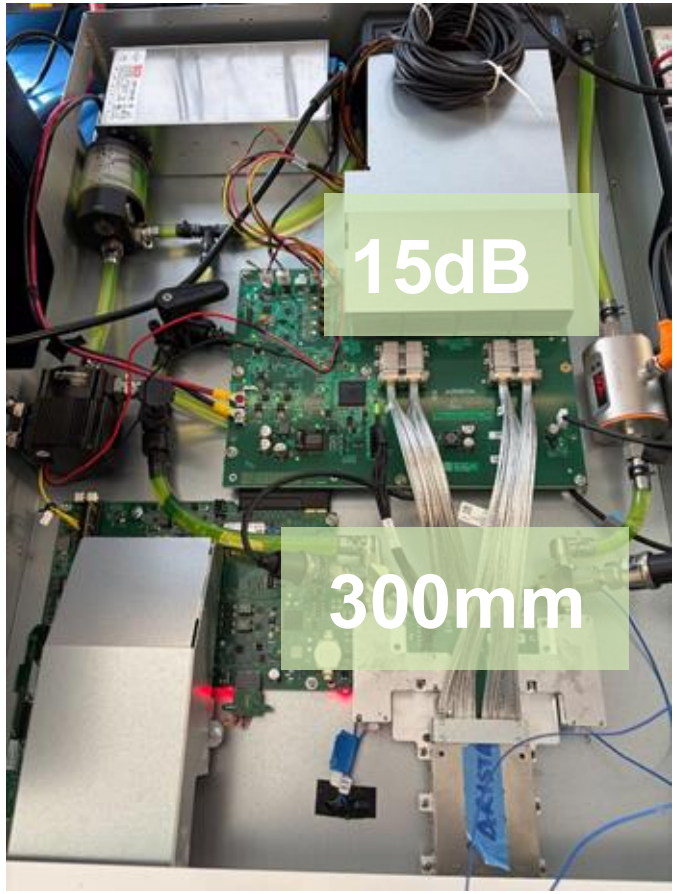
XPO Module Test Setup: End-to-End Connections

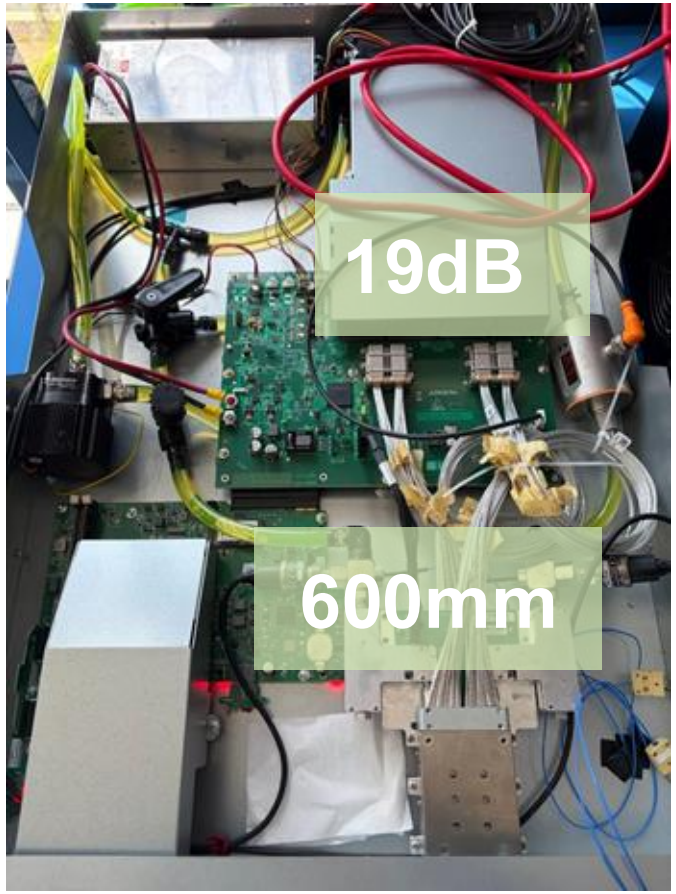


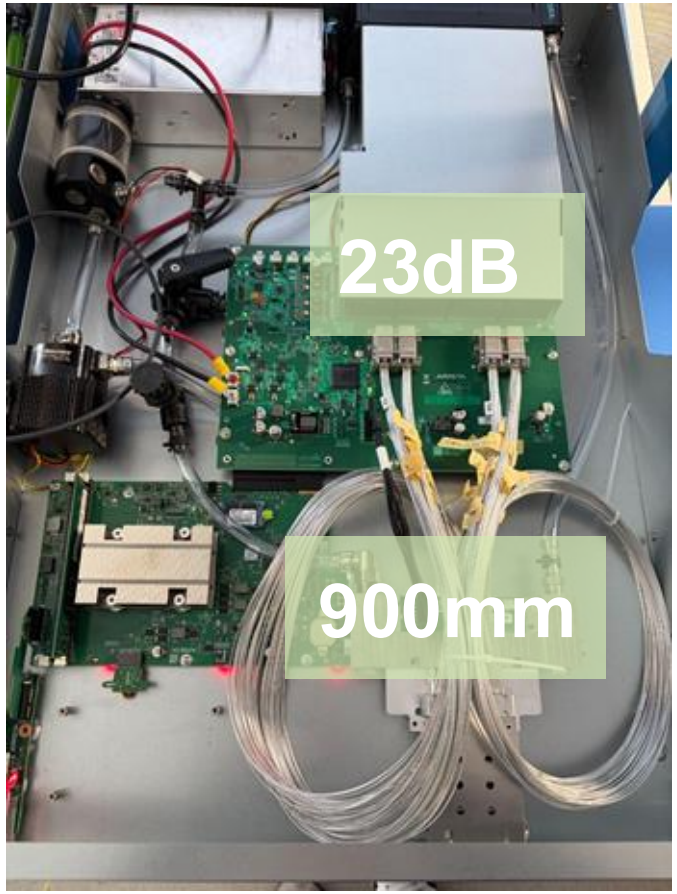
For each module version, two systems were connected, a switch on the left with fixed cable and an accelerator on the right with three flyover cable lengths. Host BER/SNR and module RX power data are collected across:

- Different flyover cable length
- W/Wo 3dB optical attenuators

XPO Module Test Setup: Accel Flyover Cable Lengths



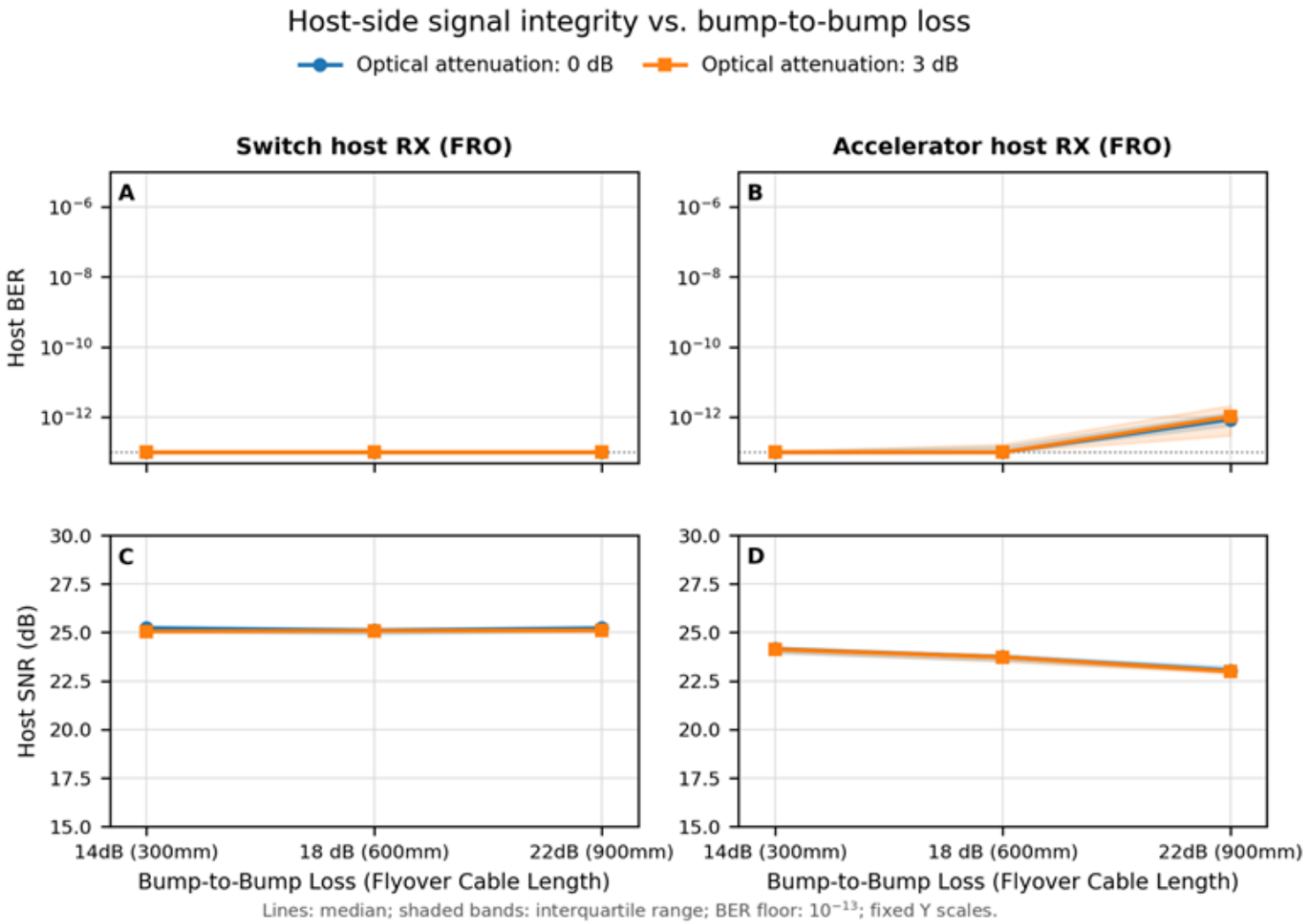




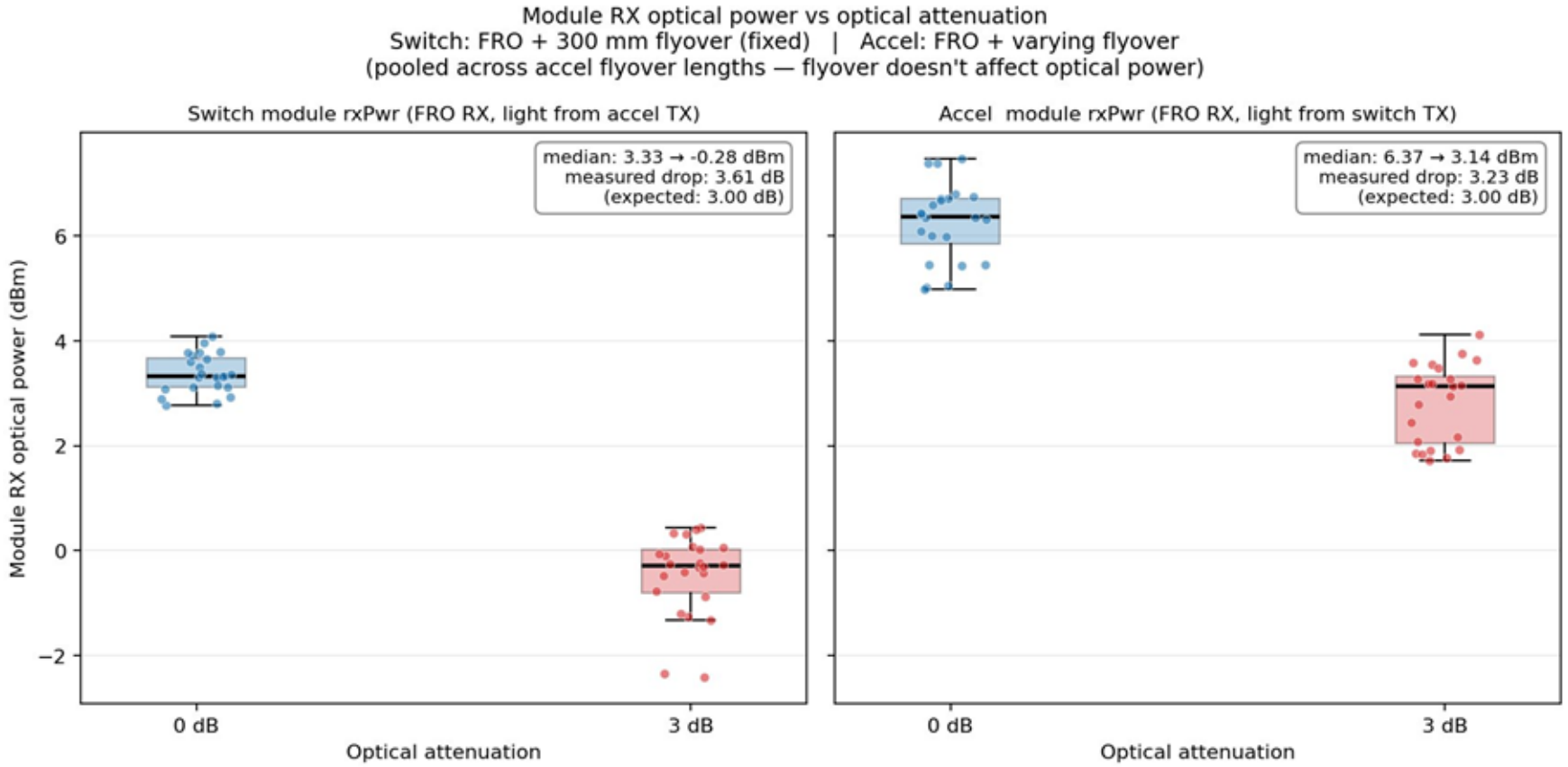
The switch side flyover cable length was fixed at 300mm (15dB bump-to-bump loss).

FRO-FRO XPO Modules: BER/SNR

- Switch-side BER remains independent of accelerator flyover cable length
- Accelerator-side BER shows minor degradation with loss
- FRO link demonstrates negligible performance differences between 0dB and 3dB loss scenarios.



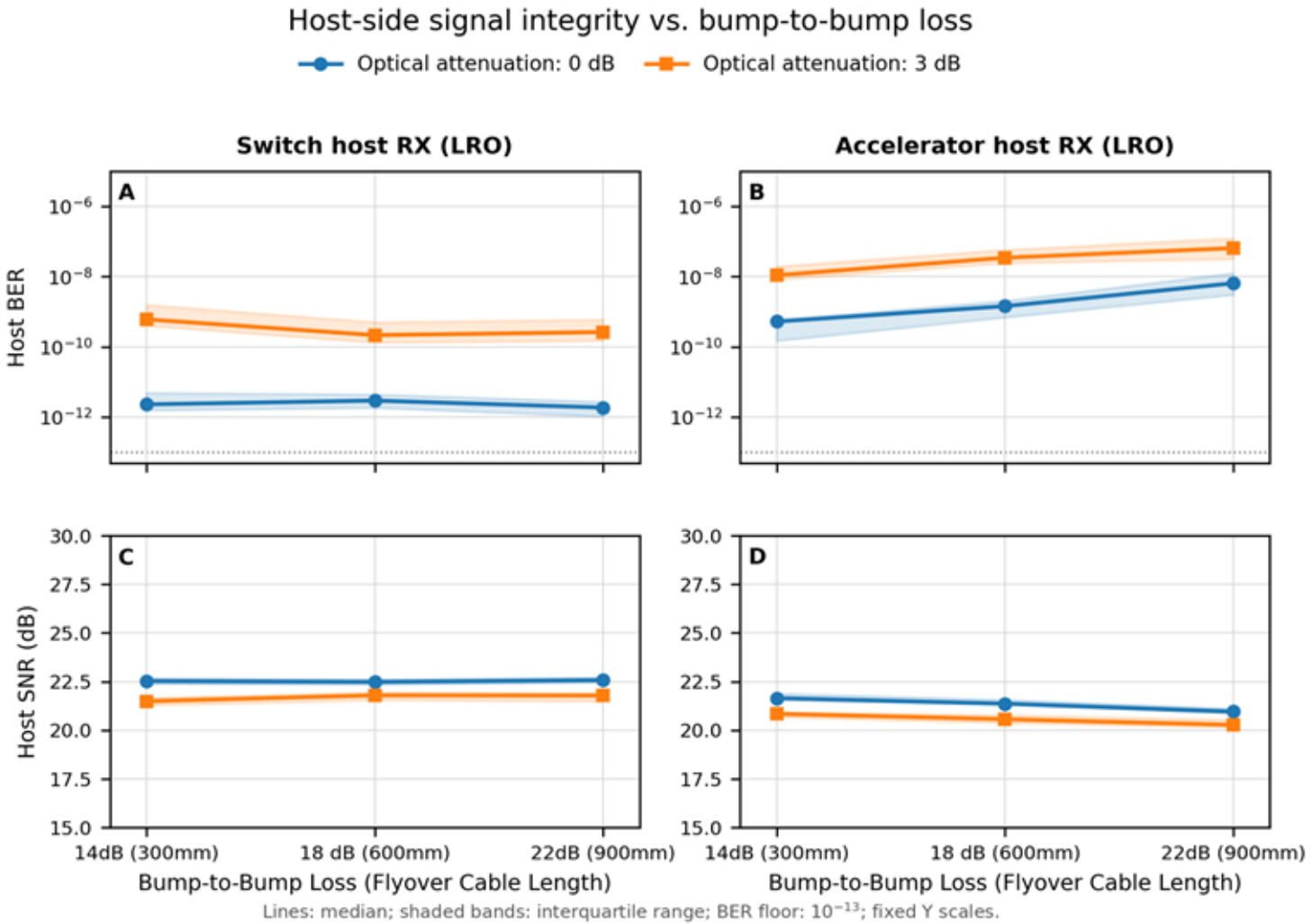
FRO-FRO XPO Modules: Rx Power



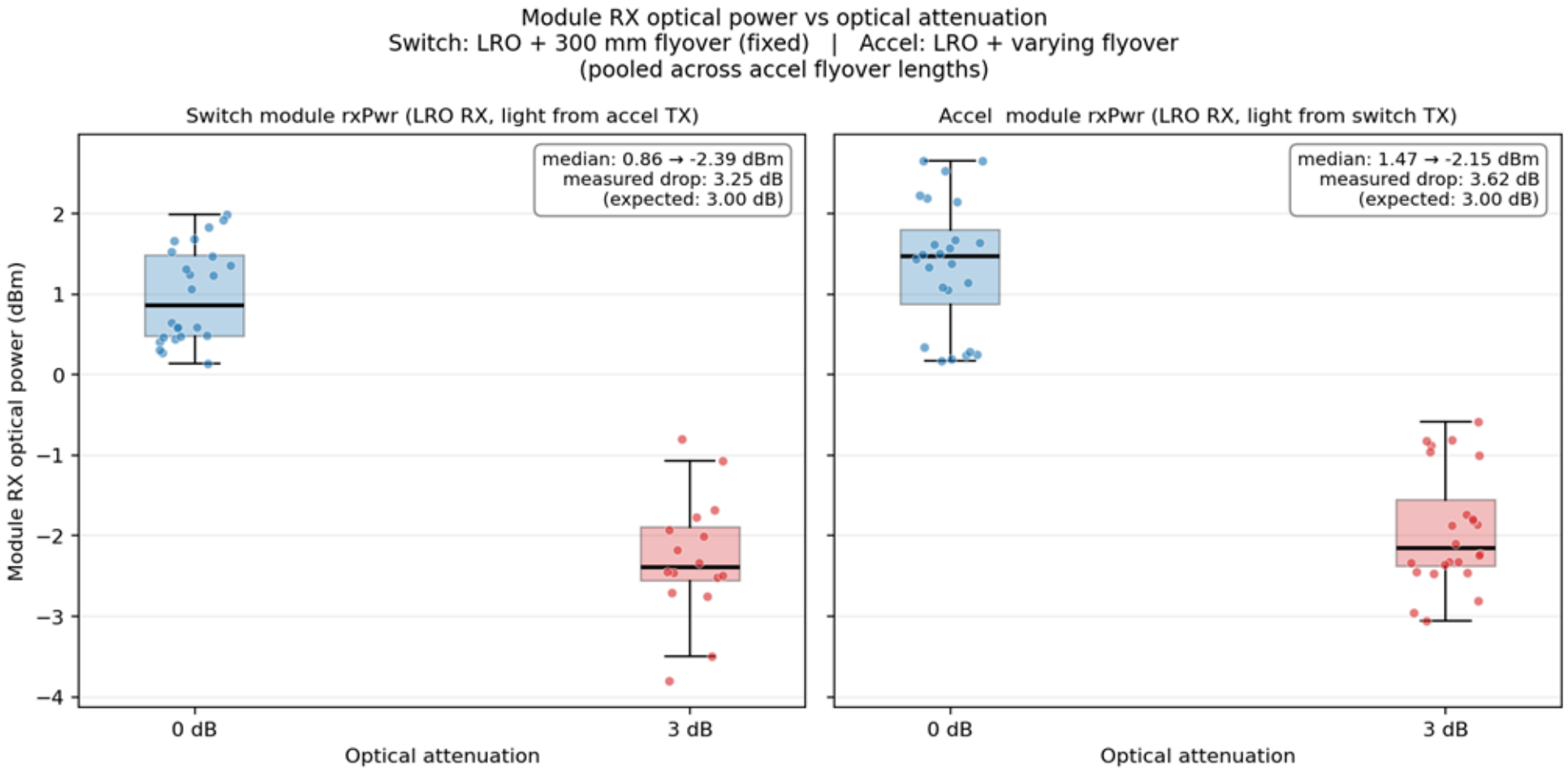
- Currently, the optical power of one module is high. As module design continues to improve, power will become more standardized.

LRO-LRO XPO Modules: BER/SNR

- Switch-side BER remains independent of accelerator flyover cable length
- Accelerator-side BER shows minor degradation with loss
- Due to the lack of Rx-side DSP, optical performance is sensitive to optical loss

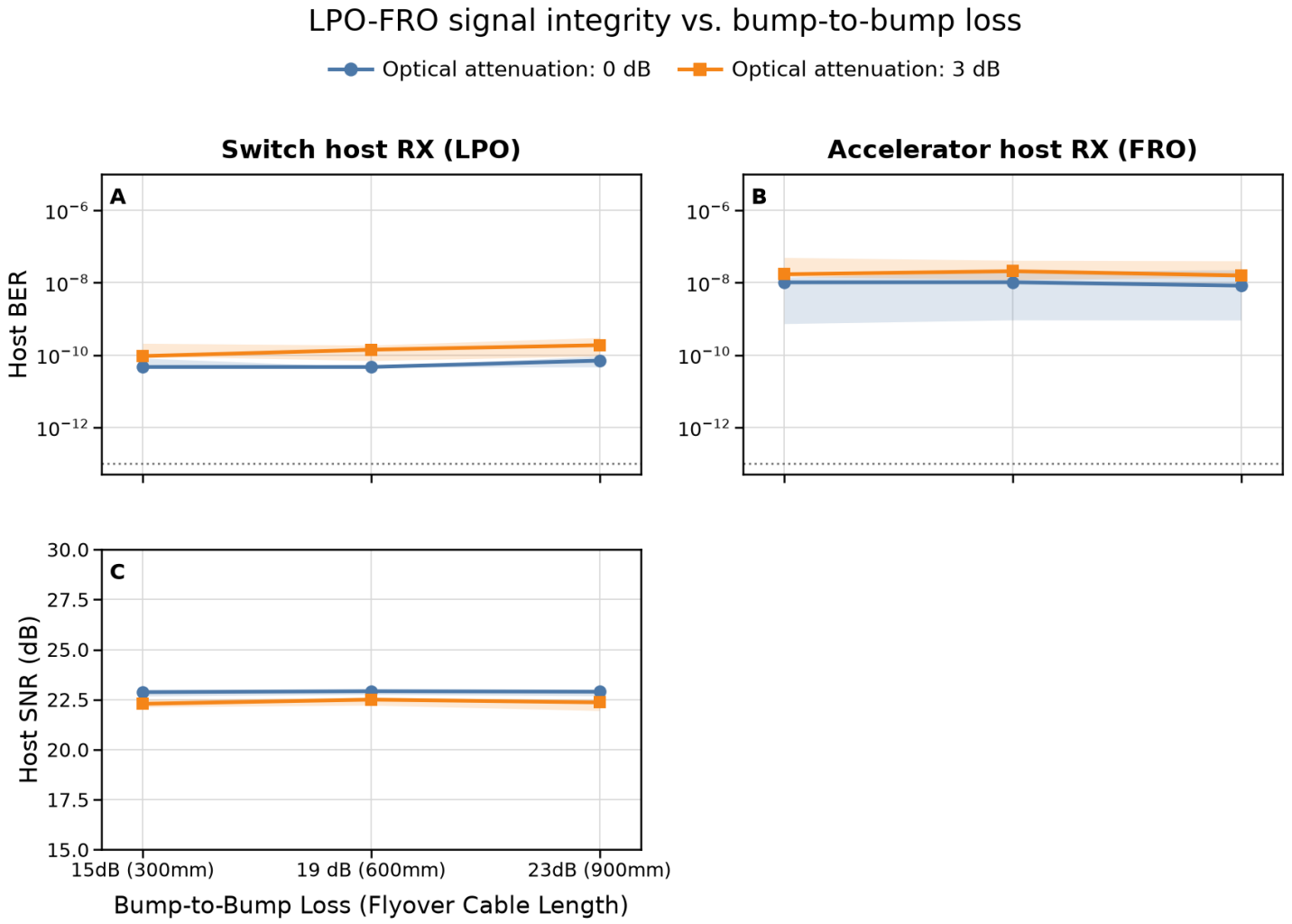


LRO-LRO XPO Modules: Rx Power



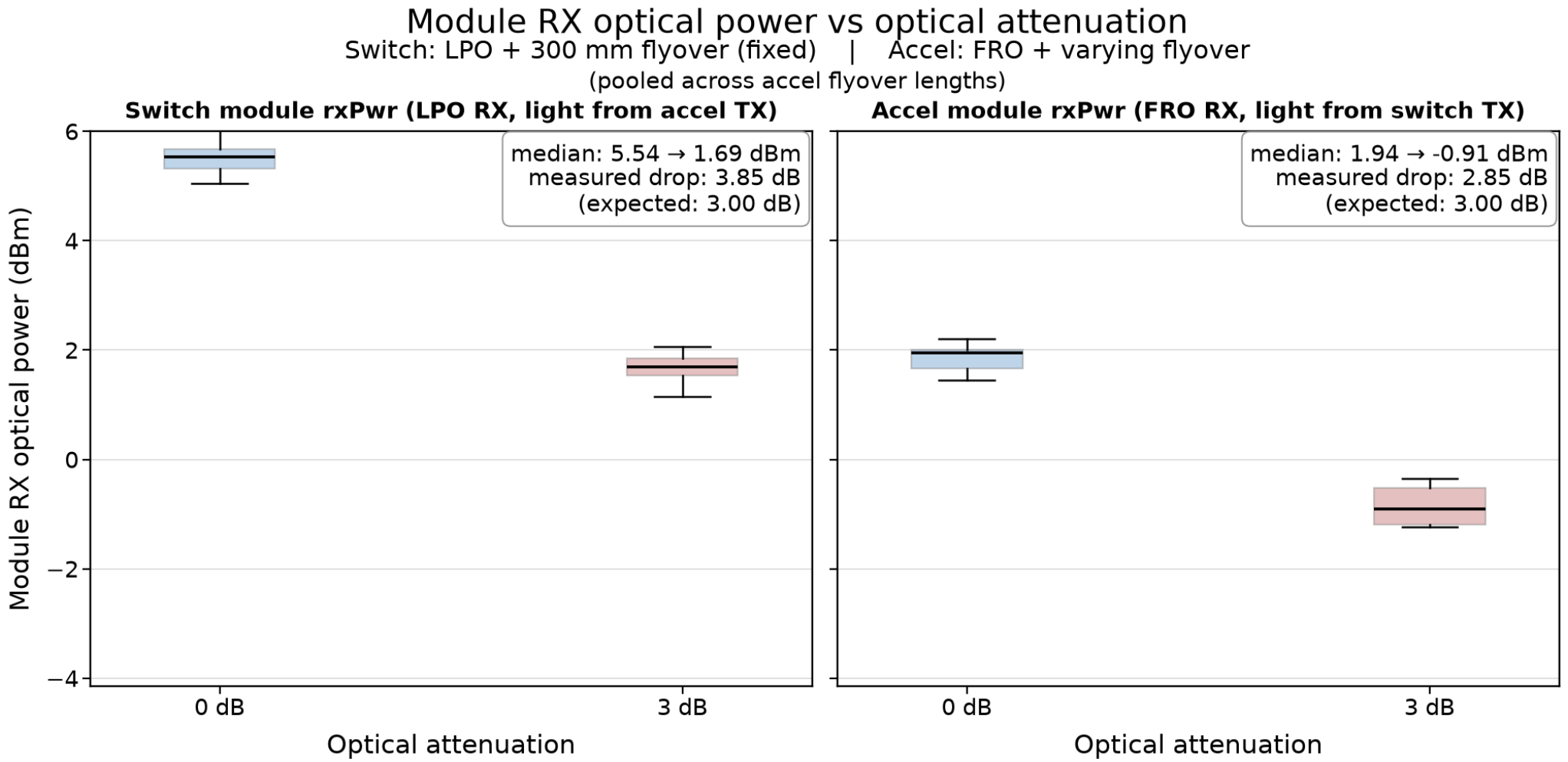
LPO-FRO XPO Modules: BER/SNR

- Switch-side BER floor stays within 10^{-10}
- Accelerator-side BER remains within acceptable 10^{-7} range, primarily constrained by LPO Tx performance
- Accelerator SNR omitted as it is not meaningful
- Link performance expected to improve with standardized power, drivers, and CTLE settings

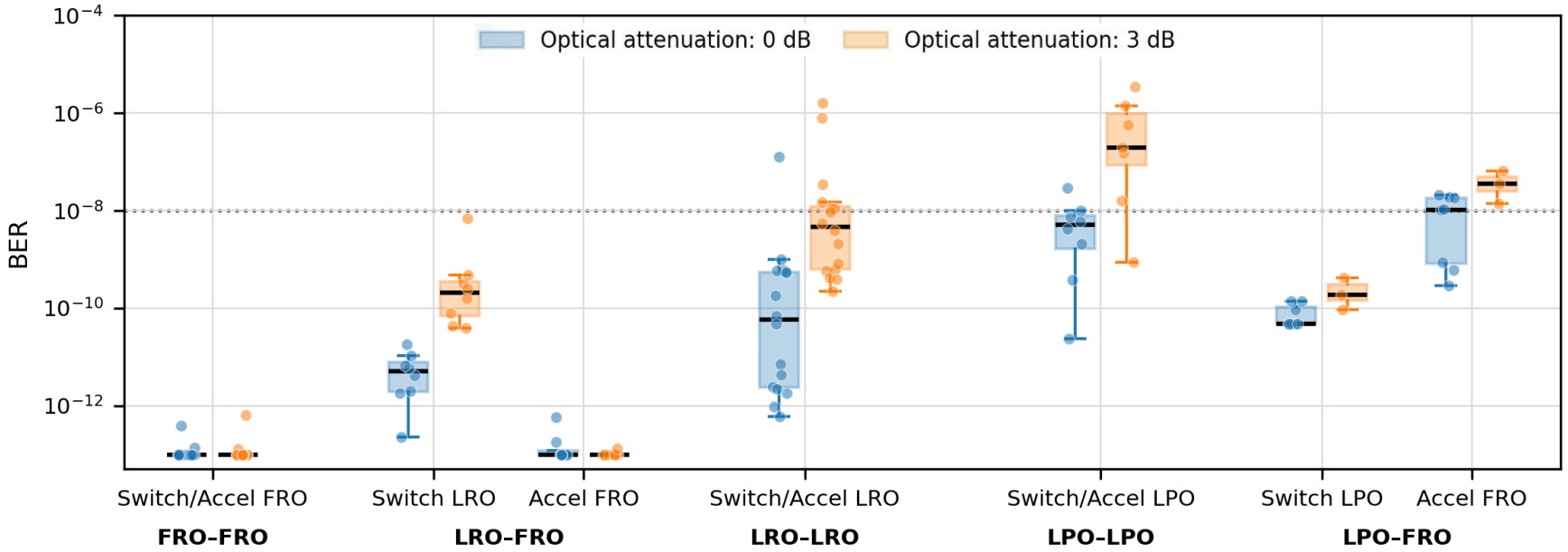


Lines: median; shaded bands: interquartile range; BER floor: 10^{-13}; fixed Y scales.

LPO-FRO XPO Modules: Rx Power



BER Comparison versus modules pairing



15 dB Bump-to-Bump Loss (300mm Flyover Cable) Module Pairing

- FRO: Provides the strongest BER margin and robust performance across the tested conditions.
- LRO: Good performance on both switch and accelerator sides, with greater sensitivity to optical attenuation.
- LPO: Demonstrates good switch-side performance; further Tx optimization and tuning can improve link margin.

Next Steps

From Demonstration to Deployment

- Expand testing to additional XPO module vendors and optical architectures
- Validate interoperability across FRO, LRO and LPO module combinations
- Optimize LPO Tx performance and establish interoperable tuning guidelines
- Progress from emulation-platform testing to switch and accelerator system-level validation

Call to Action

Build an Open and Interoperable XPO Ecosystem

- Drive multi-vendor interoperability across XPO implementations.
- Align host, module, and optical interfaces to enable plug-and-play interoperability.
- Establish common system-level test methodologies and performance metrics.
- Foster collaboration among system, silicon, and optics vendors to accelerate XPO adoption in AI scale-up networks.
- Enable new applications and innovations leveraging the XPO platform, including XPO ACC/AEC, OCI optics, coherent optics, external light sources, and emerging optical architectures.