100GEL C2M Channel Reach Options & System Design Impacts

Jane Lim, Cisco

Pirooz Tooyserkani, Cisco

Upen Reddy Kareti, Cisco

Joel Goergen, Cisco

Marco Mazzini, Cisco

5/23/2018

Supporters

- Gary Nicholl, Cisco
- Mark Nowell, Cisco
- Matt Traverso, Cisco
- David Chen, AOI
- Kohichi Tamura, Oclaro
- Rajesh Radhamoha, Maxlinear
- Hassaan Aslam, Maxlinear
- Phil Sun, Credo
- Takeshi Nishimura, Yamaichi
- Chris DiMinico, MC Communications/PHY-SI LLC
- Rachad Samaha, Multilane

100GEL C2M Channel Requirements (Cisco Proposal in Rosemont Meeting)

- Target ball to ball IL ≤ 12dB @ 26.56GHz :
  - To avoid using heavy receiver inside the optical modules → lower module power consumption
  - Make 100GEL C2M link budget work with 100GEL CR → twinax Cu cables support upto 2 m (see Goergen_100GEL_02_0318)
    - **ball to ball 30dB @26.56GHz, bump to bump 36dB @26.56GHz**
- Overall IL target should allow reasonable trace length with minimum 5” to reduce # of retimers at front ports
- Include present & next generation packaging and PCB technologies

What Have Been Decided in Rosemont Meeting

- Define a single-lane 100 Gb/s PHY for operation over electrical backplanes supporting an insertion loss ≤ **28** dB at **26.56** GHz.

- Define a single-lane 100 Gb/s PHY for operation over twin-axial copper cables with lengths up to at least **2** m.

100GEL C2M & CR Link Budget Original Proposal

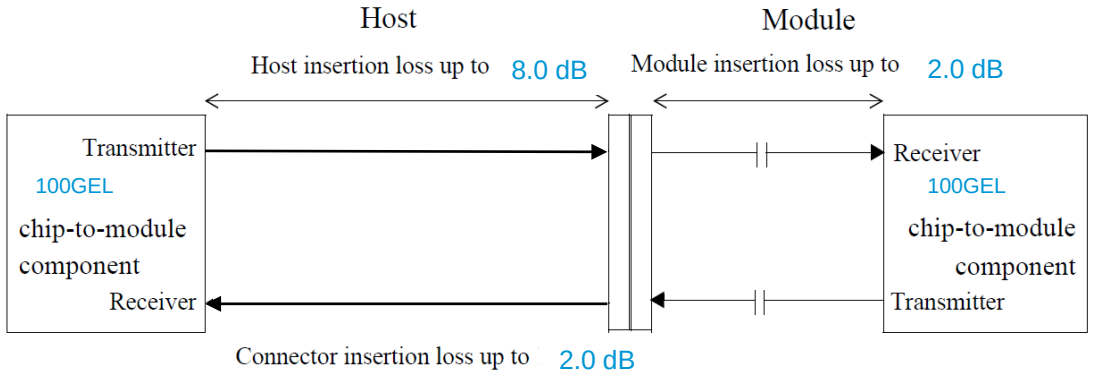


Figure 1: 100GEL C2M insertion loss budget at 26.56 GHz

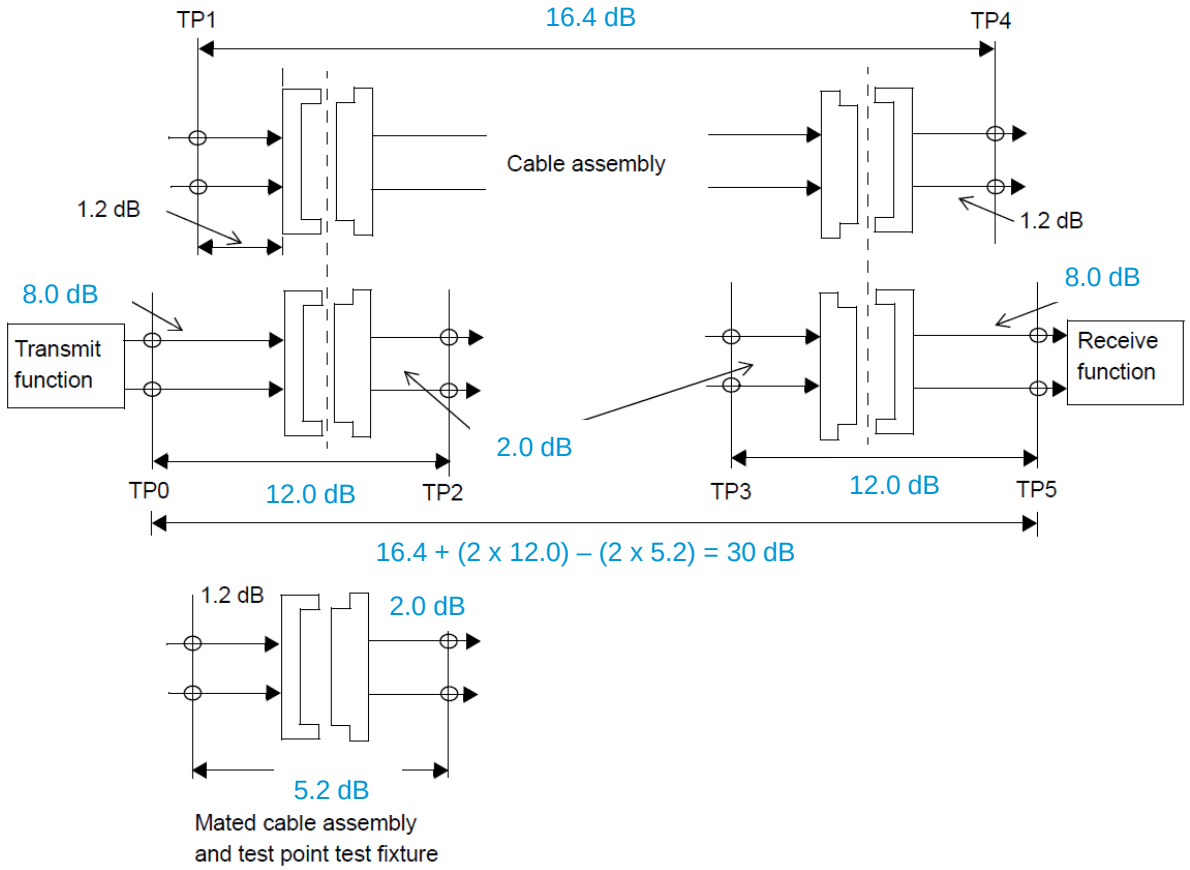


Figure 2: 100GEL CR 30dB insertion loss budget at 26.56 GHz

100GEL C2M & 100GEL CR Link Budget Adjusted Proposal

- Tighten host PCB budget, from 8.0 dB to 7.5dB
- Tighten connector only loss, from 2.0 dB to 1.5 dB
- Loosen Module PCB or HCB loss, from 2.0dB to 2.5dB

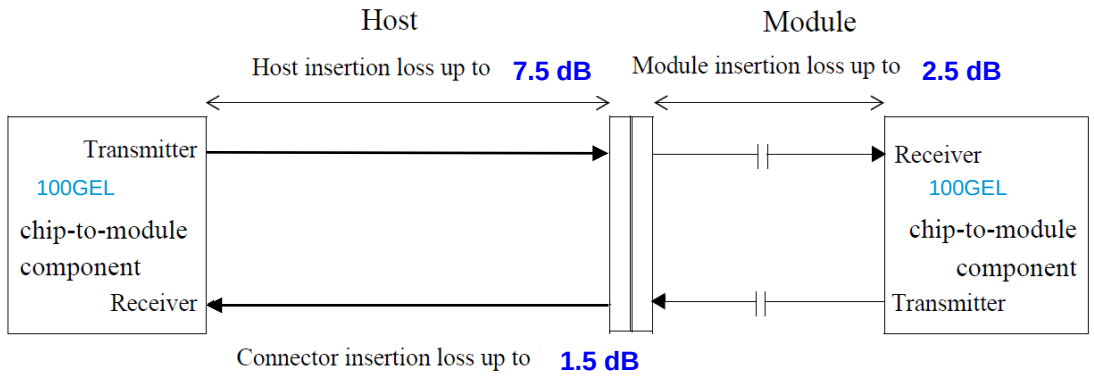


Figure 1: 100GEL C2M **11.5dB** insertion loss budget at 26.56 GHz

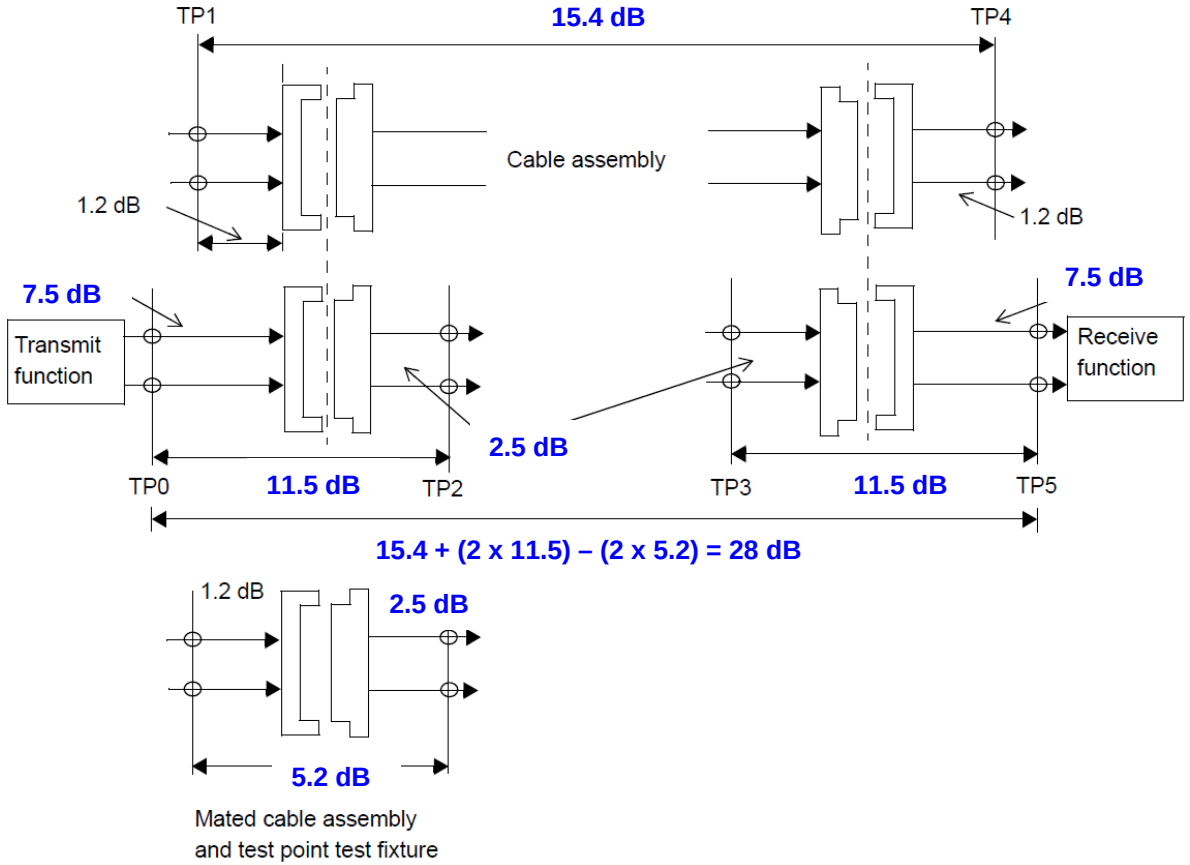


Figure 2: 100GEL CR **28dB** insertion loss budget at 26.56 GHz

Why 7.5dB is Required for Host PCB Budget ?

- With Meg-7N material, 4.5mil trace IL at HT is measured to be 1.24 dB/in at 28 GHz
- For each front port channel there will be 2 set of vias (at host ASIC BGA footprint & at I/O connector footprint) with stripline routing
  - Footprint via with 7.9mil drill & 130mil thick stackup is simulated to be 0.68 dB at 28 GHz
- Total host PCB budget = (5 x 1.24 + 2 x 0.68) = **7.56 dB**

100GEL C2M Channel Options

- 3 Options in consideration
  - Option 1: 11-12dB VSR-like reach target
    - Symmetrical link budget on switch side (12dB) & NIC side (12dB)
    - Same port can support 2m DAC
  - Option 2: 14-16dB SR-like reach target
    - Asymmetrical link budget - switch side (15dB) & NIC side (9dB)
      - In switch-to-switch case, this is equiv. to Option III in “ghiasi_100GEL_adhoc_01_050918” which require separate switch host design for 2m DAC support
  - Option 3: 18-20dB MR-like reach target
    - Extended reach for optical interface
    - No DAC support

Option 1: 11-12 dB VSR-like Reach Target (1)

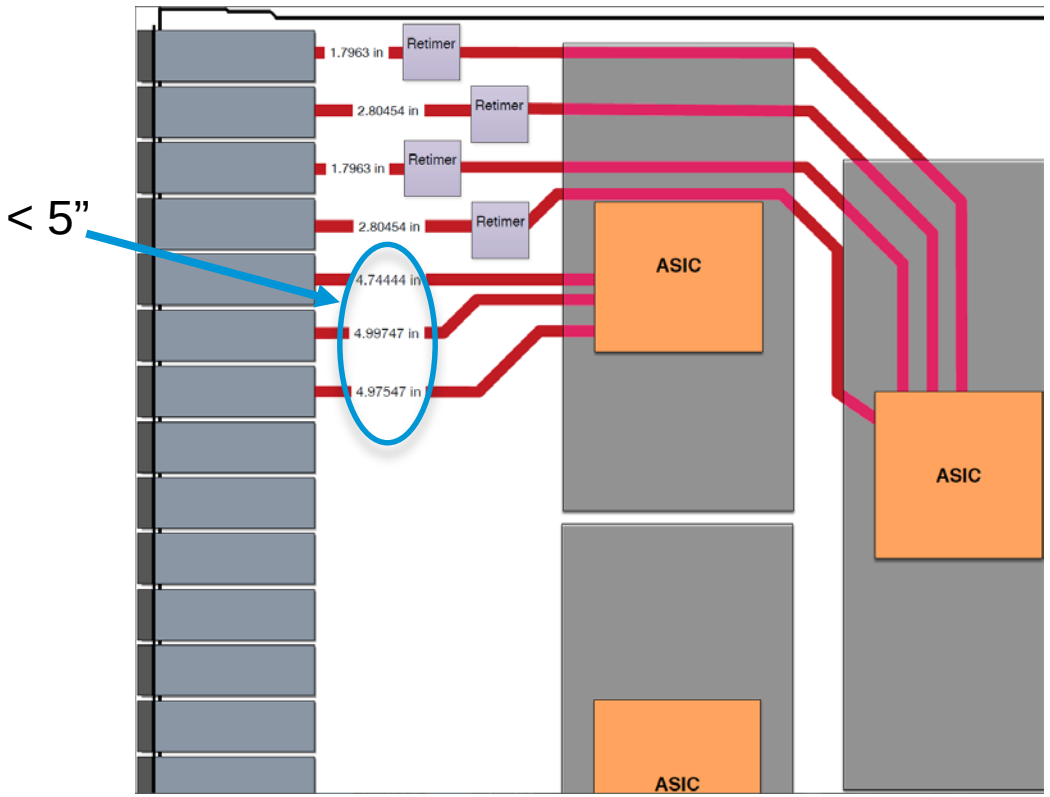
- Pros:
  - Can support both Cu and Optics in the same port (Universal port), similar to what are accustomed to today
  - Benefits lower power optical module design (simple serdes)
    - 100 – 300 mW/ch power saving compared to MR-like serdes
  - May need to add PHY device anyhow on every port to support MACsec or other features
  - Offers flexibility in system designs, i.e. for channel > 5”, can use retimers or intra-box cables (see next slide), belly to belly module connector, etc

- Cons:
  - Potential impact on system cooling when using intra-box cables
  - System cost increase using retimer or intra-box cables
  - Require chiplet or innovative packaging technique to reduce package loss

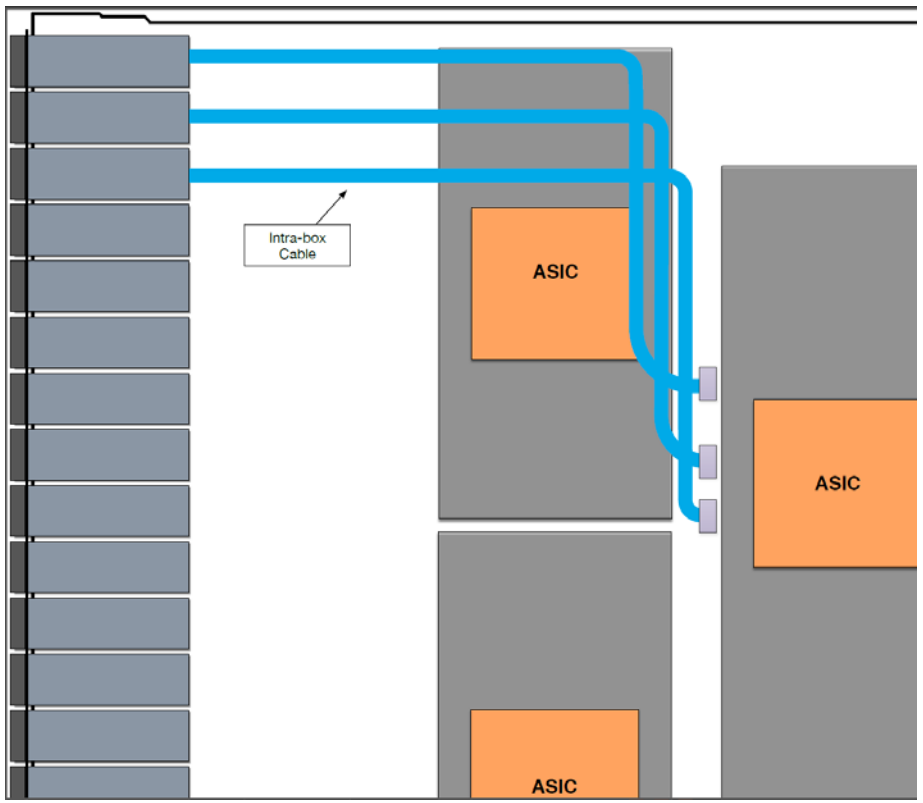
Option 1: 11-12 dB VSR-like Reach Target (2)

- Universal port design on multi-ASIC line-card to support both Optics and 2m DAC
- The diagram depicts the actual placement and routed trace length

Edge ports use retimers



Edge ports use intra-box cables



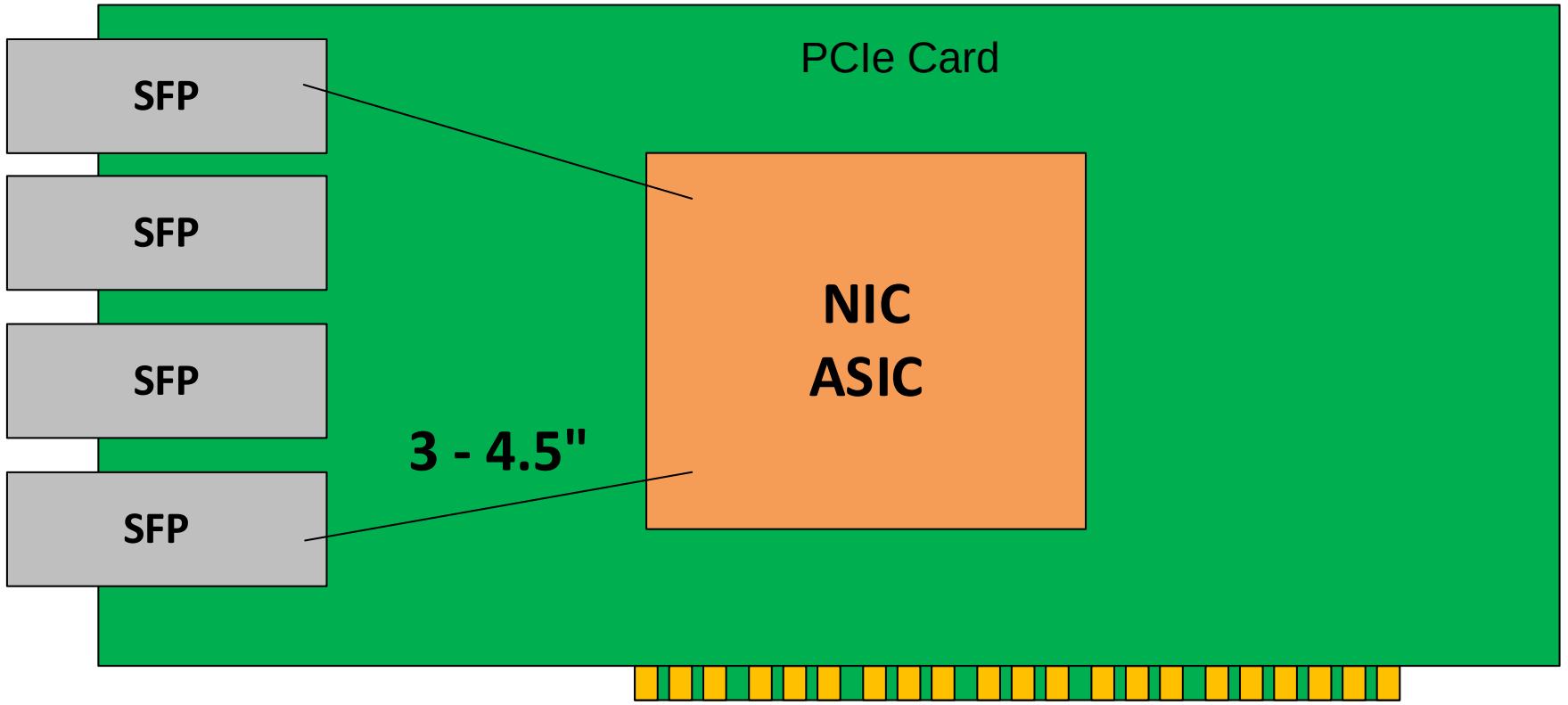
Option 2: 14-16 dB SR-like Reach Target (1)

- Pros:
  - Allow longer host PCB trace for Optics (upto 8”), much fewer retimers or intra-box cables
  - Less impact on system cooling
  - System design cost advantage (less retimers)
  - All the flexibilities in option 1 plus a choice for the system designers to optimize the non-DAC case

- Cons:
  - Switch to switch 2m DAC support doesn’t work
  - More complex Rx design in module CDR
    - Serdes power penalty (need serdes suppliers to provide incremental power increase)
    - Most Serdes vendors are not considering to design for this SR-like Serdes at present time
  - More expensive for rack server NIC card, may need to add retimer for >3” trace length (see next slide)

Option 2: 14-16 dB SR-like Reach Target (2)

- Rack Server PCIe NIC Card:
  - Ethernet SFP Port typical trace lengths of 3 – 4.5” based on today NIC card implementations. Potential trace length reduction to 4” depends on ASIC pin-out and board placement



Option 3: 18-20 dB MR-like Reach Target

- Pros:
  - Allow even longer host PCB trace (upto 11”) for Optics
  - No retimers or intra-box cables needed (which is the main goal), lower system cost
  - Less impact on system cooling without intra-box cables

- Cons:
  - **No support for 2m DAC (against the adopted objective !)**
  - Very complex Rx design in module CDR → Power penalty
    - Significant power increase compared to VSR (upto 1.2W added serdes power for a 400G module)
  - Impact on the design of certain PMDs (exceeding module power envelope)
  - Serdes may require link training (in-band or out-band)

Summary

- Our preference for C2M reach is Option 1 with 11-12dB which aligns with 2m DAC objective, and gives simple product consumption for DAC cables (switch to switch, switch to server, server to server)
- We are still studying Serdes design feasibility and module design implication for Option 2, but have concern in product consumption and impact to the server design
- Option 3 with MR Serdes is not a viable solution

Thank You !

Backup Slides

400GAUI-8 C2M IL Budget

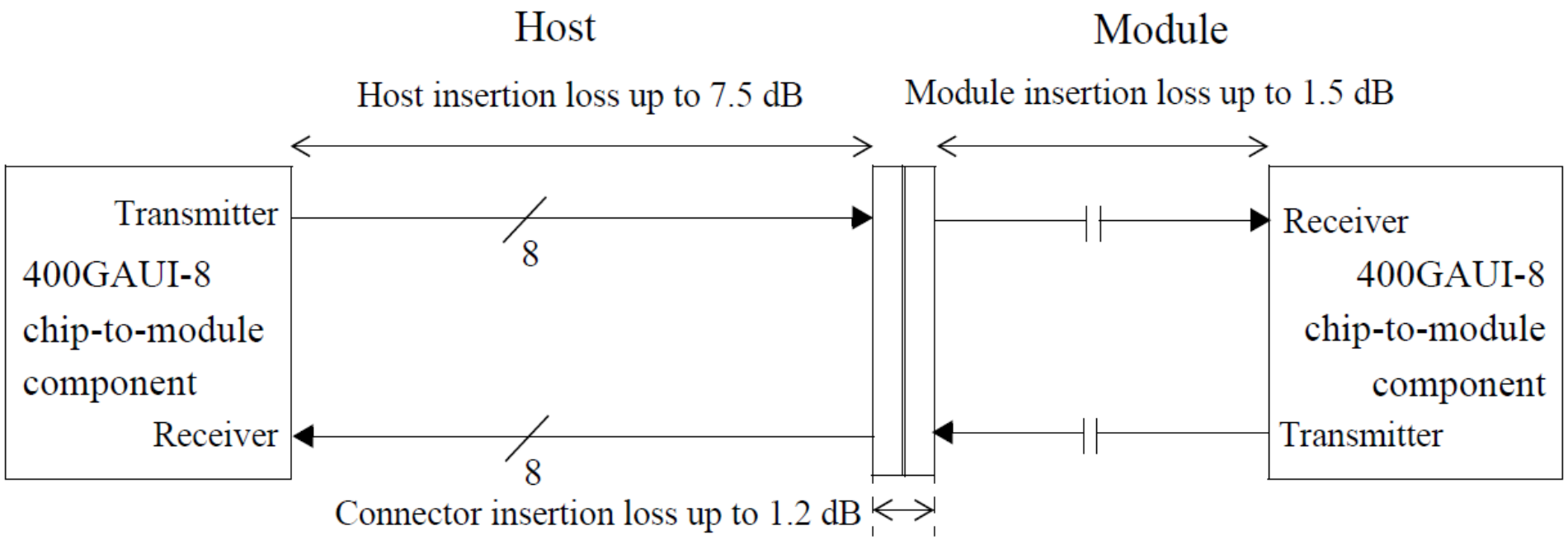
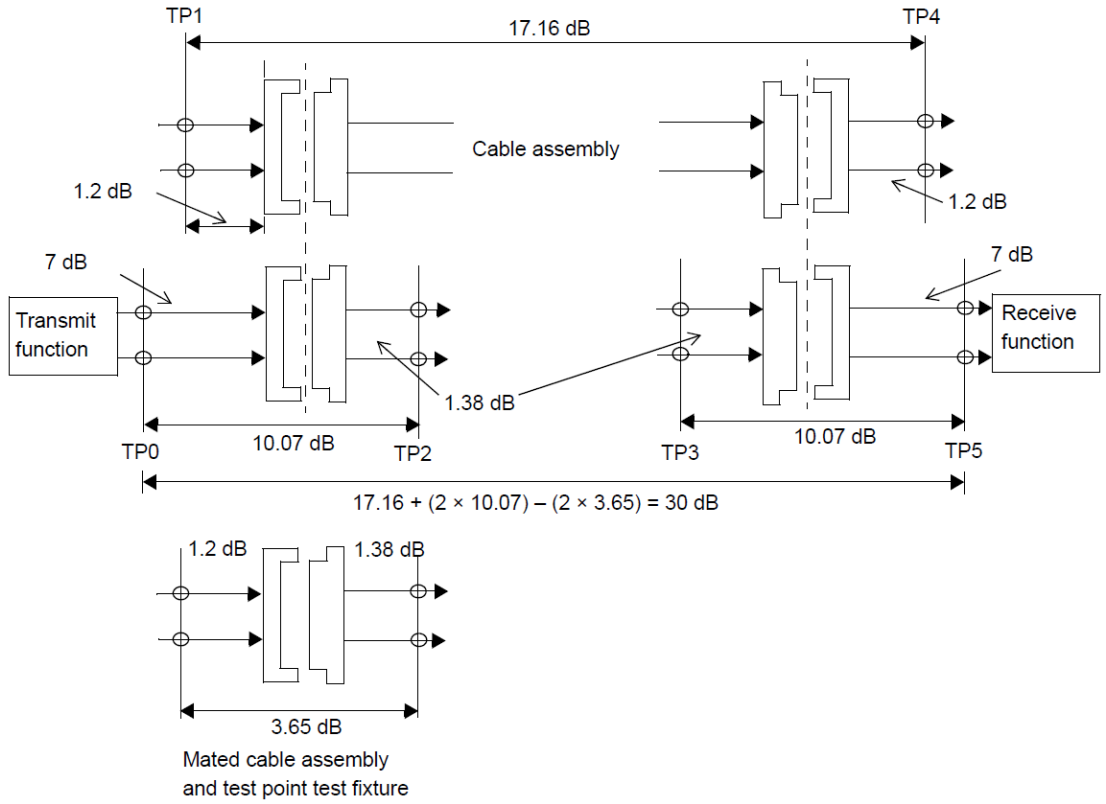


Figure 120E–3—400GAUI-8 chip-to-module insertion loss budget at 13.28 GHz

Total channel link budget (ball to ball) = 10.2dB at 13.28GHz

Source: IEEE802.3bs

200GBASE-CR4 Channel IL Budget



NOTE—The connector insertion loss is 1.07 dB for the mated test fixture. The host connector is allocated 0.62 dB of additional margin.

Figure 136A–1—30 dB channel insertion loss budget at 13.28 GHz

Source: IEEE802.3cd