Physics Driven Paths to Scalable Accelerator Fabrics

PSU Harrisburg: *19th Central PA Symposium of SI/PI*

Andrew Josephson

April 17, 2026

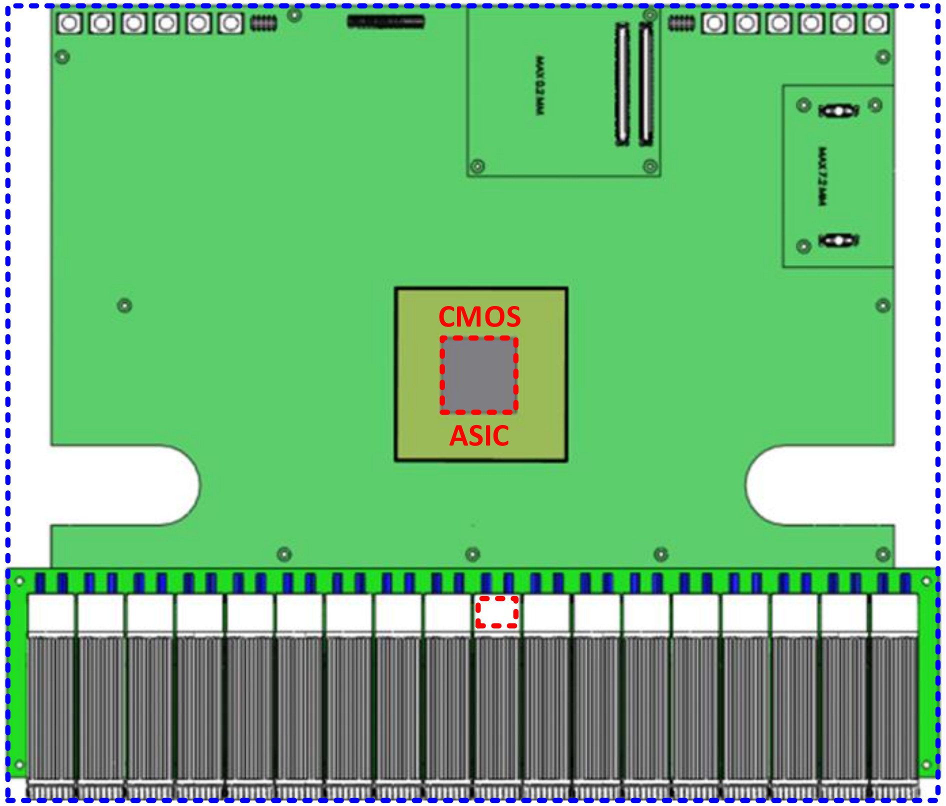
INNOVATIVE TECHNOLOGIES • SUDDEN SERVICE • GLOBAL REACH

Agenda

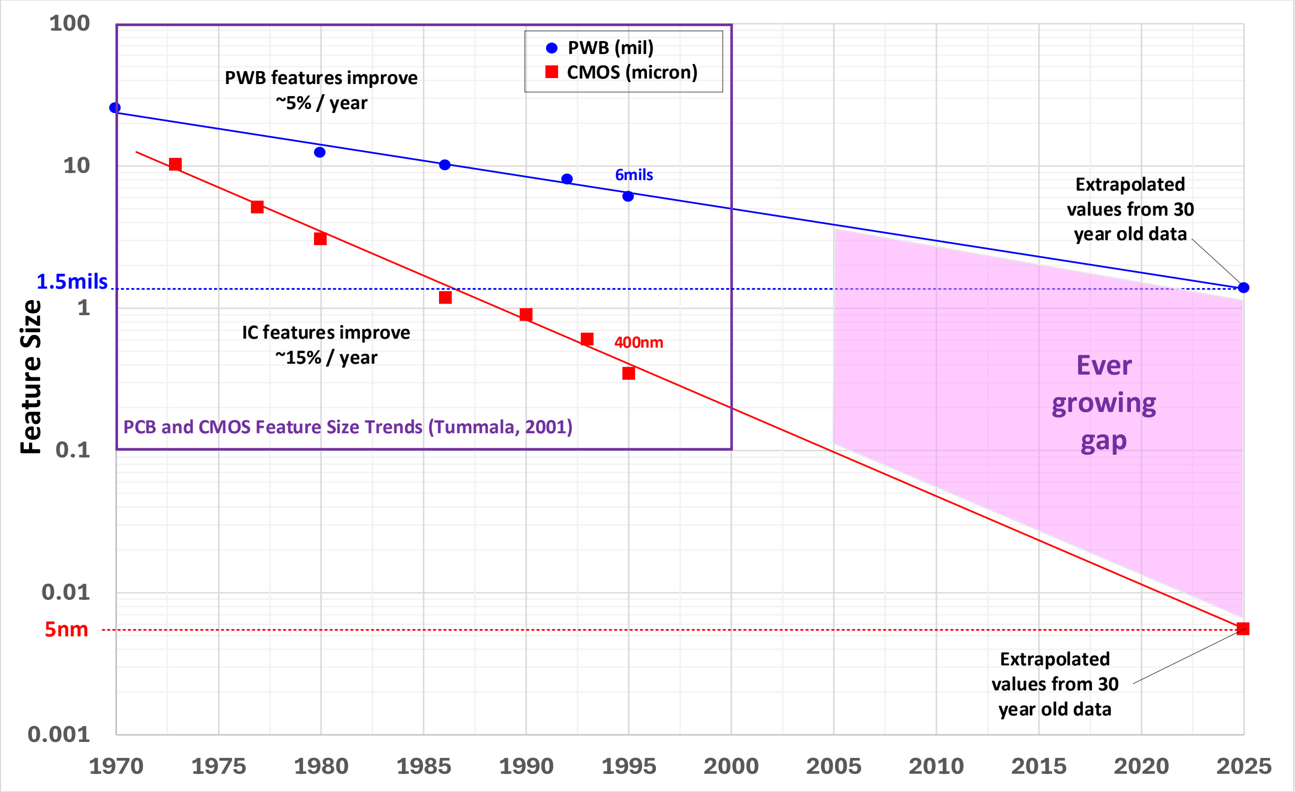
- Interconnect Feature Size
  - Shrinking wavelength scale, UI and interconnect feature size
- Background Physics and Resonances
  - Quarter Wave Stubs, Emitters, Slots, Higher Order Modes, Fiber Weave
- Apply radiation physics concepts to PCB interconnect scale features at ~80GHz for 400G signaling spectrums of interest.
  - Ask PCB features radiation questions instead of SI interconnect ones
  - Listen for textbook emitter answers.
  - Overlay radiation limits on PCB feature size limiting PCB SI performance (loss/inch)
- PCB reach constrained topologies
- Compute Cluster BW Metrics

Interconnect Feature Size

PCB vs. CMOS Scaling

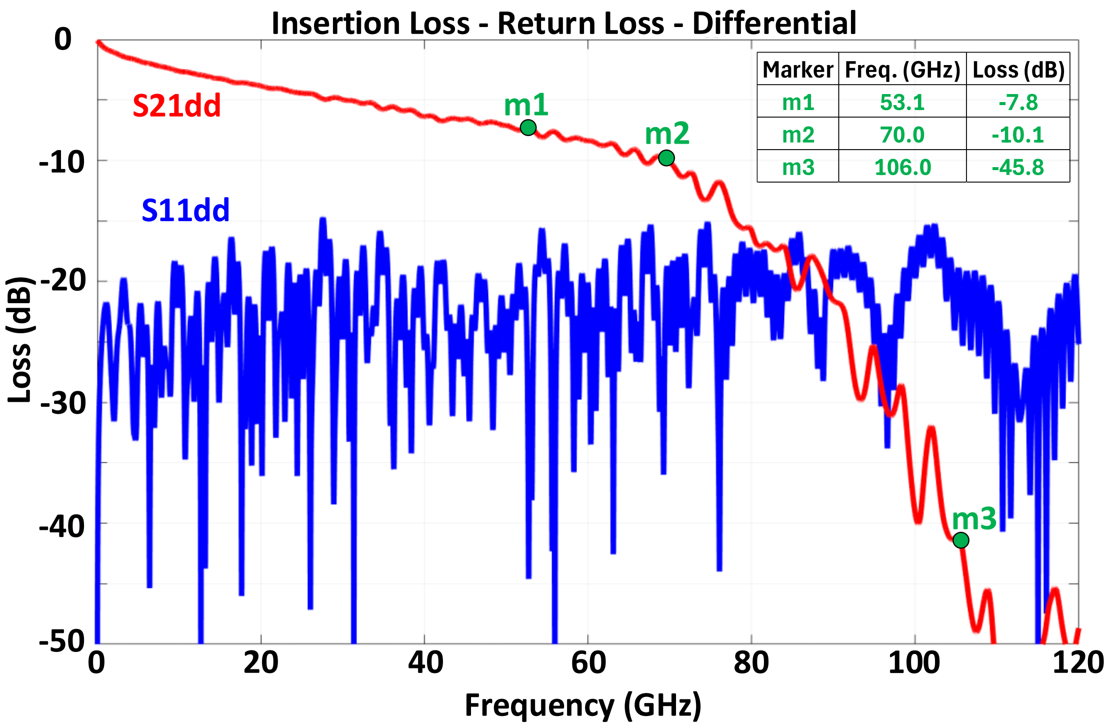


We want this concept to work from 1Gbps to 400Gbps

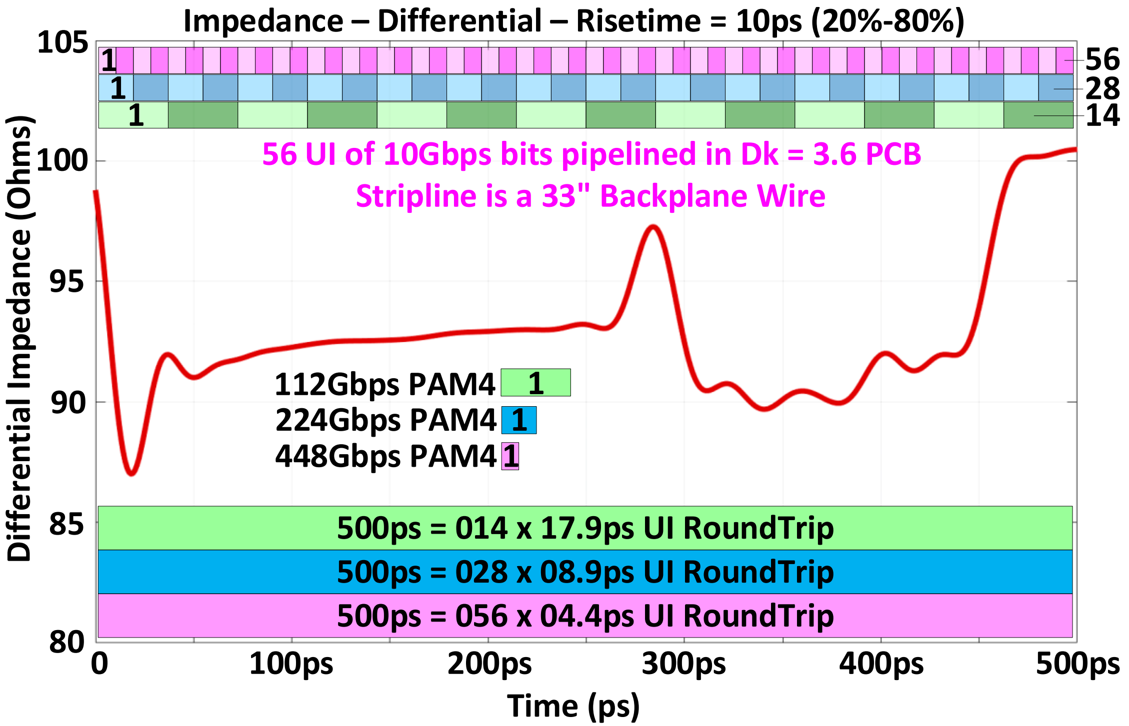


- PCB and CMOS feature scale trends from 1970-1995 Extrapolated to 2025 suggest we should be using 40um PCB traces by 4nm.
  - Why are SI engineers stuck on 30 year old PCB features?

Shrinking UI


Insertion Loss - Return Loss - Differential

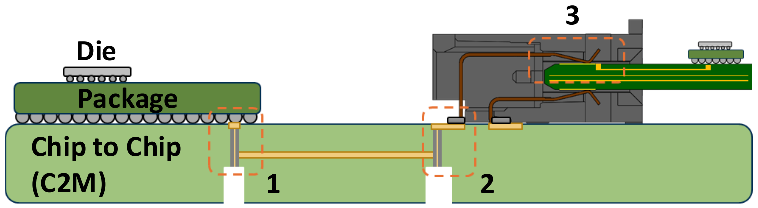
Marker	Freq. (GHz)	Loss (dB)
m1	53.1	-7.8
m2	70.0	-10.1
m3	106.0	-45.8


Impedance – Differential – Risetime = 10ps (20%-80%)

56 UI of 10Gbps bits pipelined in Dk = 3.6 PCB Stripline is a 33" Backplane Wire

112Gbps PAM4 1
224Gbps PAM4 1
448Gbps PAM4 1

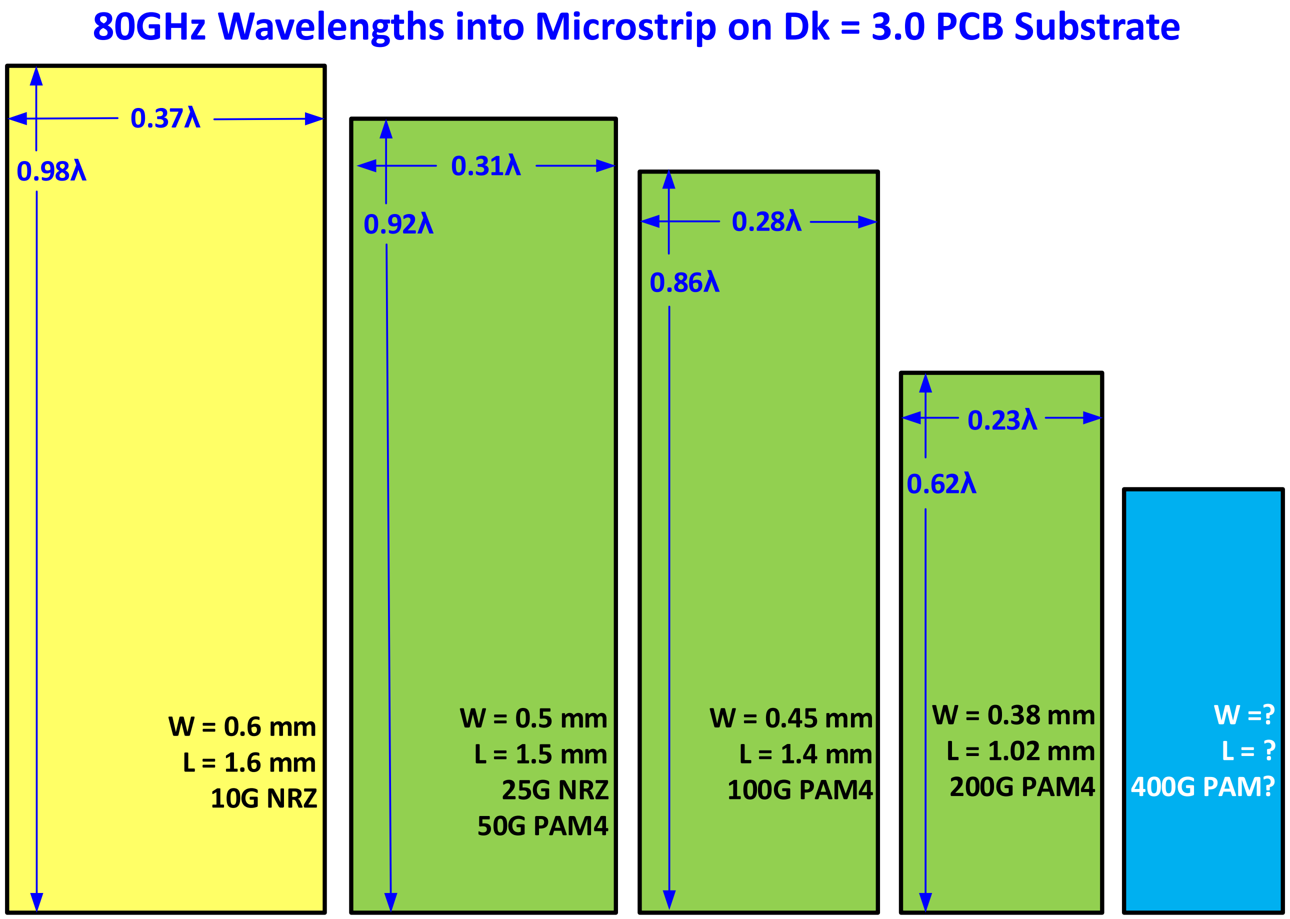
500ps = 014 x 17.9ps UI RoundTrip
500ps = 028 x 08.9ps UI RoundTrip
500ps = 056 x 04.4ps UI RoundTrip

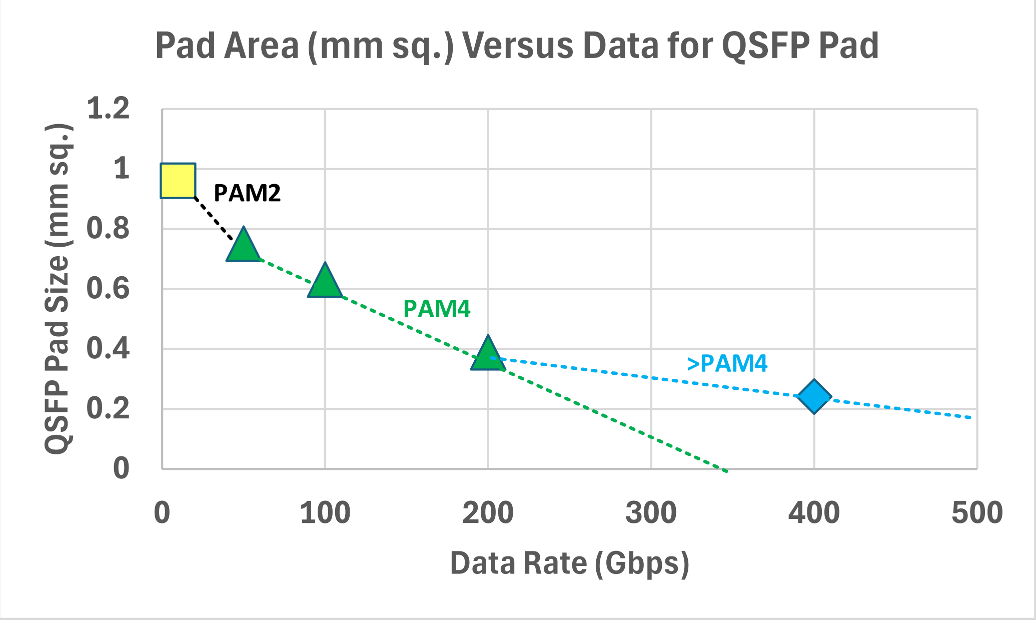


3 Reflective Problem Regions

- At 448Gbps PAM4, 1 UI = 4.4ps
  - Physically short C2M topologies of inches in length have pipelined bit counts on the order of yesterday's rack scale backplanes.
- 1.00mm of interconnect in Dk=3.0 dielectric has a halfwave electrical length at 86.6GHz.

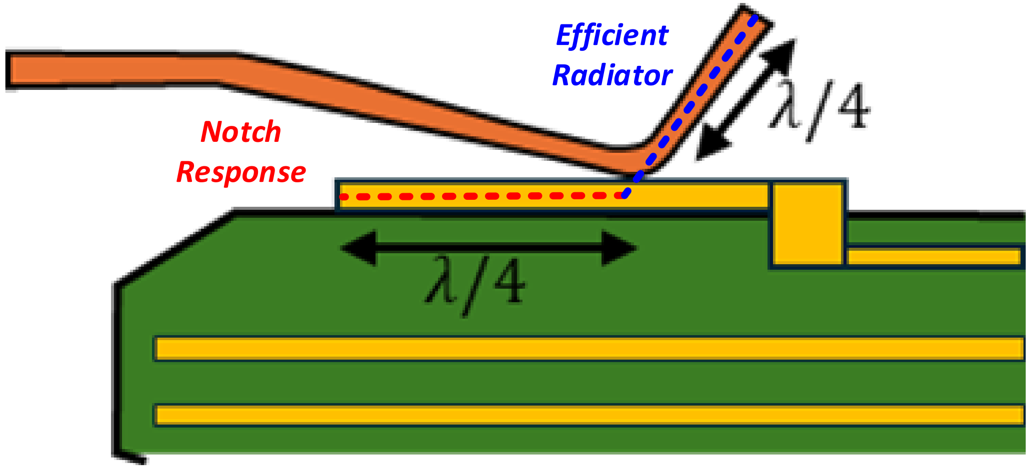
QSFP Pad Size vs. Data Rate

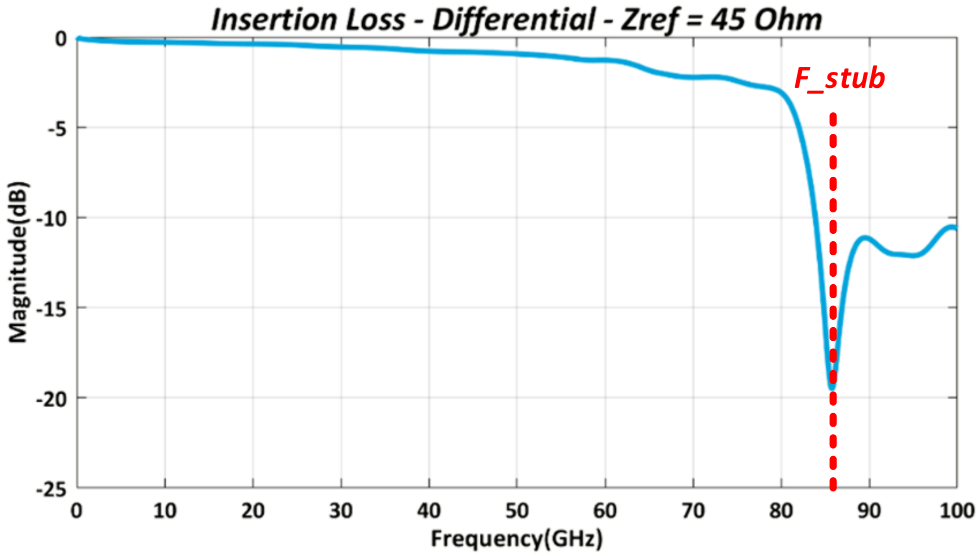




- Simple scaling of the QSFP Pad to data rate suggests higher order modulation is required if edge card contact pitch and geometries are not physically shrunk.

Quarter Wave Stubs



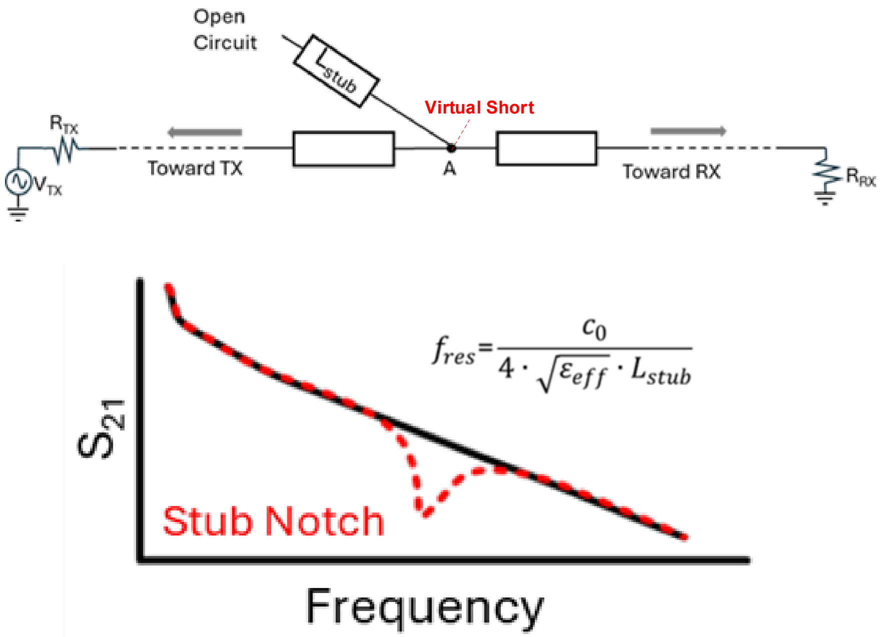


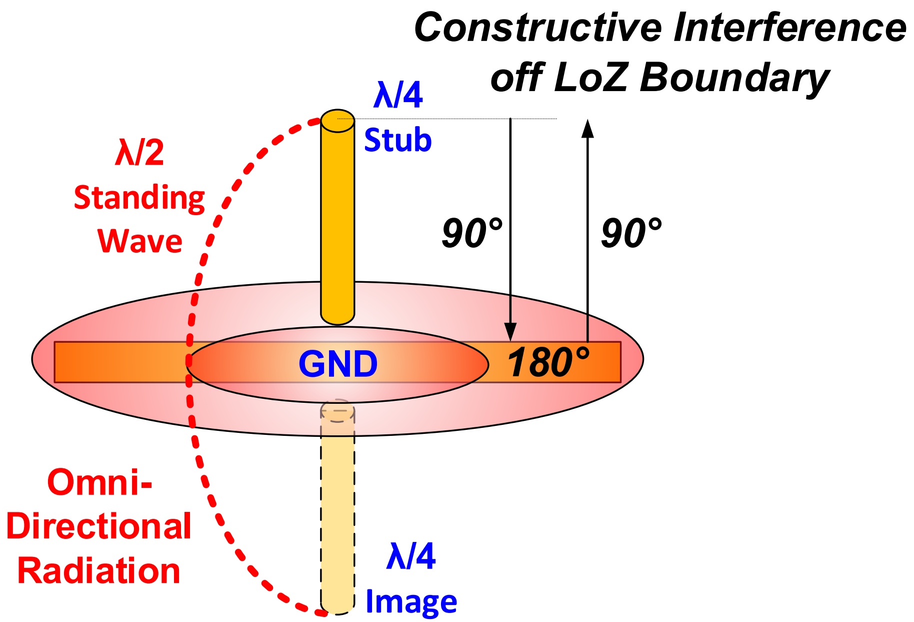
- Edge card connector beams require PCB pad and metal beam stubs to overcome mechanical tolerances and provide wipe.
  - These stubs limit achievable rolloff BW and create standing wave emitters that increase radiation and ultimately crosstalk

- Quarter wave stub brick wall from QSFP spring fingertip and PCB pad at ***F_stub***
- Shorter '*blue stub*' is occurring to the right of ***F_stub***

Background Physics and Resonances

Backround Physics





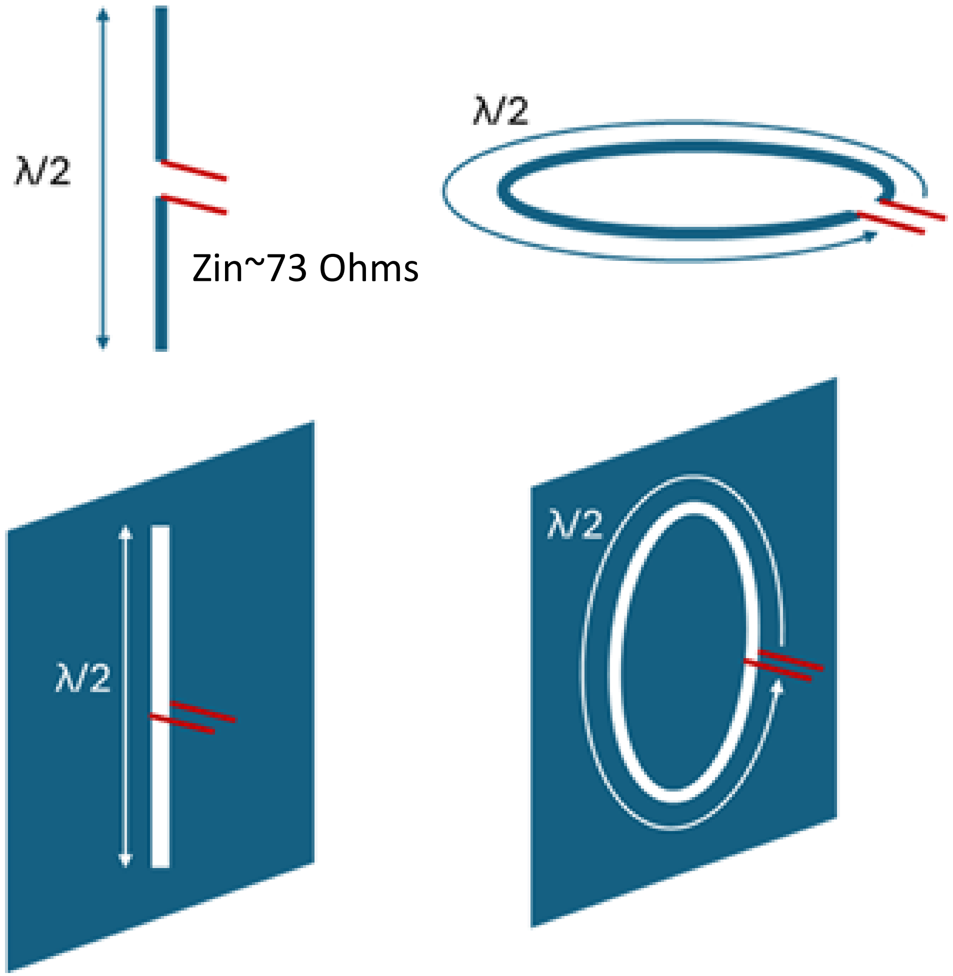
- Parasitic shunt interconnect becoming electrically limiting
  - – We still have no equalization techniques for dead shorts (or virtual ones)

- Parasitic shunt interconnect becoming efficient radiators
  - – Leaky interconnect losing SNDR to environment (BW density)

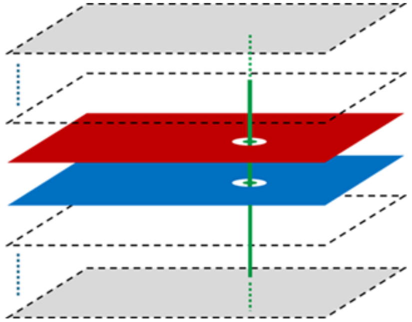
Babinet’s Principle

- Relates radiation patterns and input impedances between positive and negative feature emitters
  - Positive Feature = Wire or Stub
  - Negative Feature = Slot or Gap

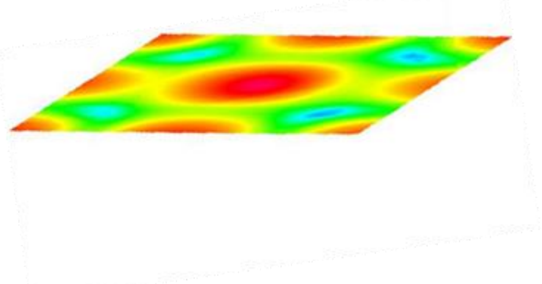
- Slot aperture impedances are normally high compared to signaling Zo
  - Encapsulating them inside PCB insulators and filling them with interconnect pads further lowers their impedance creating more efficient coupling to/from low Zo interconnect.

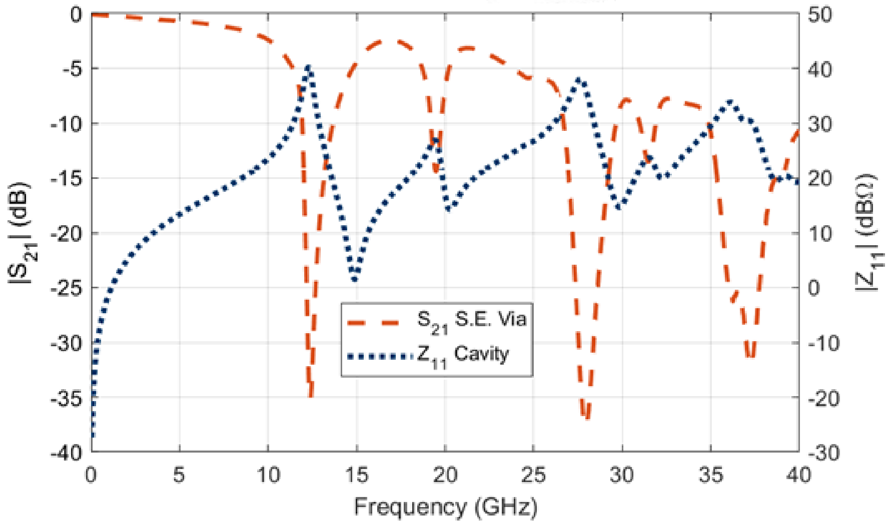


Exciting Planar Reference Plane Resonances



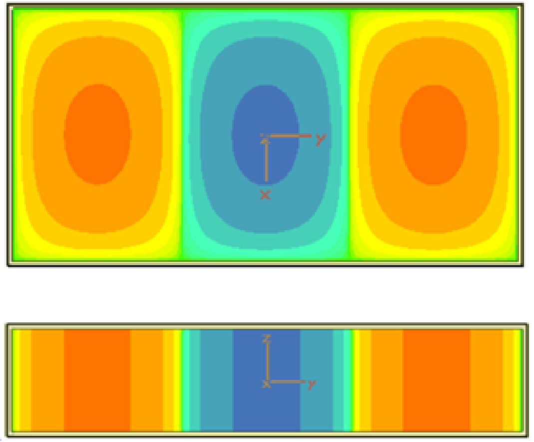
$$f_{TMmn} = \frac{c_0}{2\pi\sqrt{\varepsilon_r}}\sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2}$$



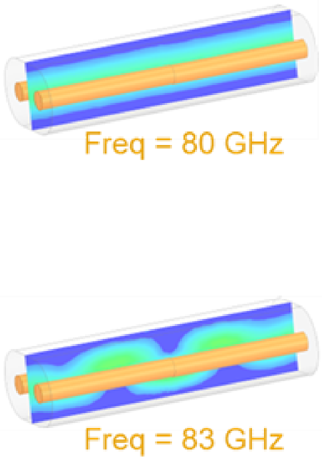


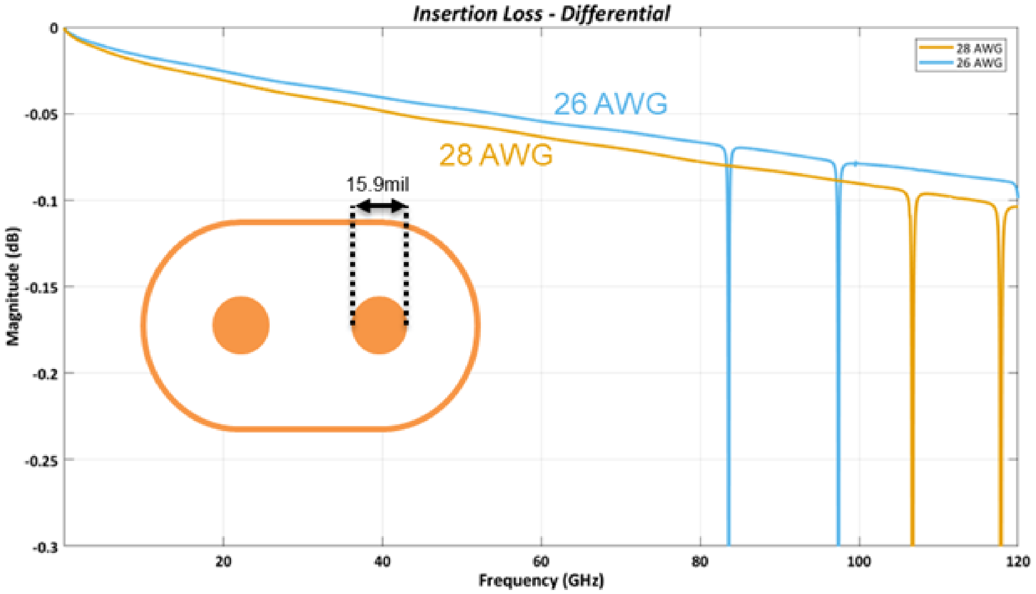
- As planar reference plane extents exceed wavelength dimensions, nonzero voltage (ground bounce) can establish standing wave phenomena
  - Presents reactive impedance in the return path (virtual opens)
  - Can establish efficient radiation conditions from the standing waves increasing crosstalk and degrading SNDR

Exciting Higher Order Modes



$$f_{TM-TE,mnp} = \frac{c_0}{2\pi\sqrt{\varepsilon_r}}\sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 + \left(\frac{p\pi}{d}\right)^2}$$

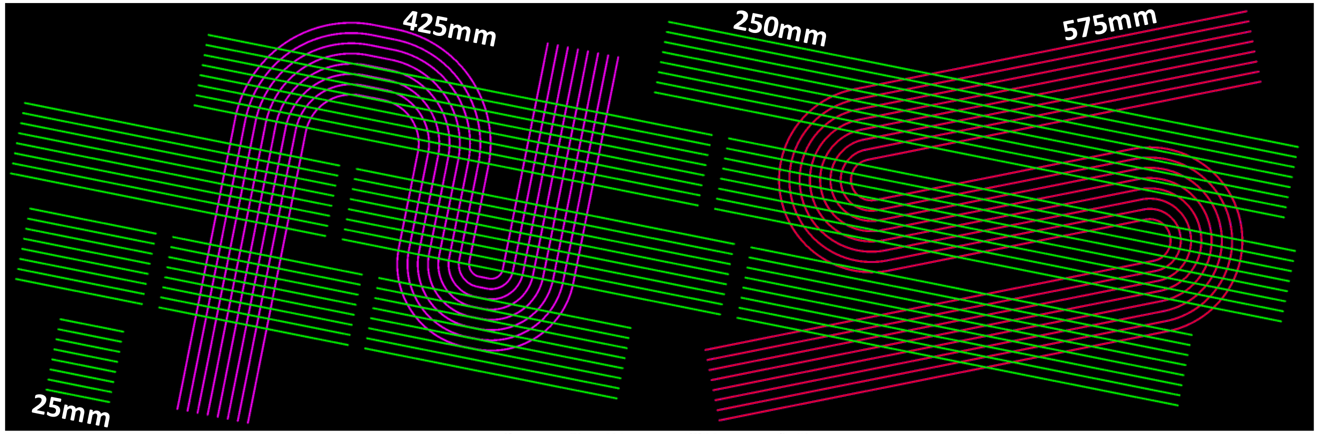


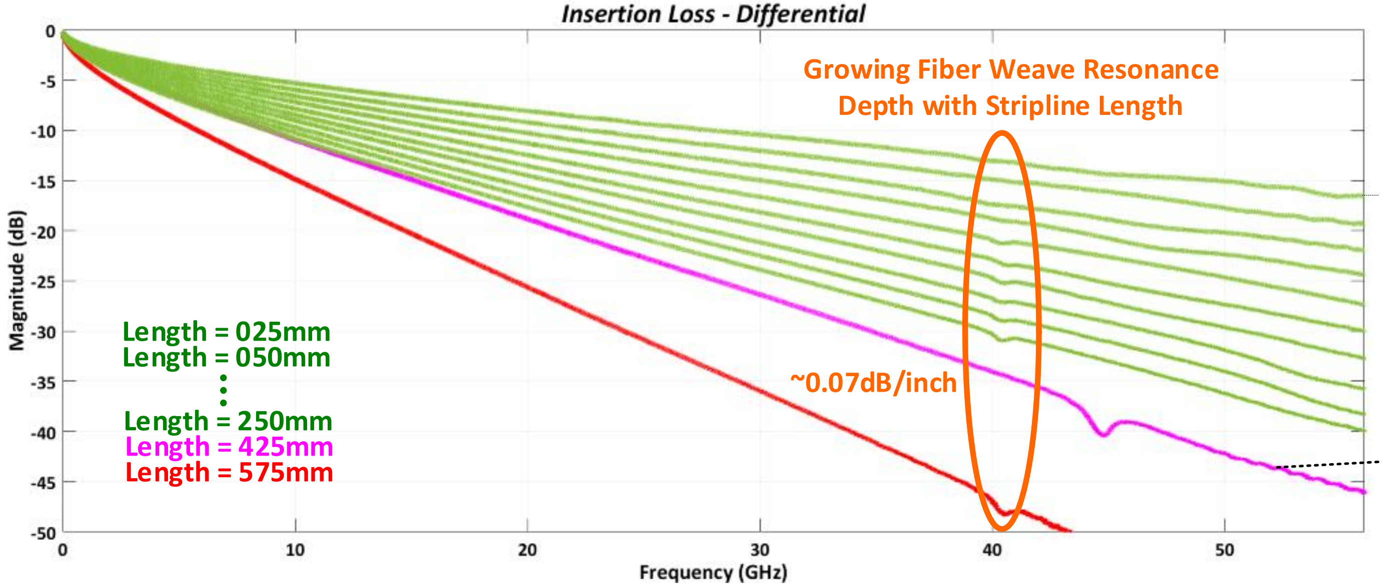


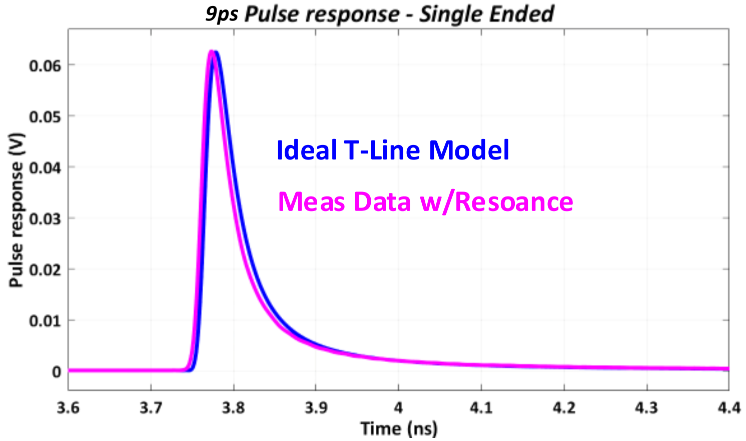
- Higher order modes are possible when half wavelength dimensions fit within a waveguide wall
- Targeted transmission mechanism in higher order waveguides
- Evanescent reflectionless loss mechanism in coax and TwinAx

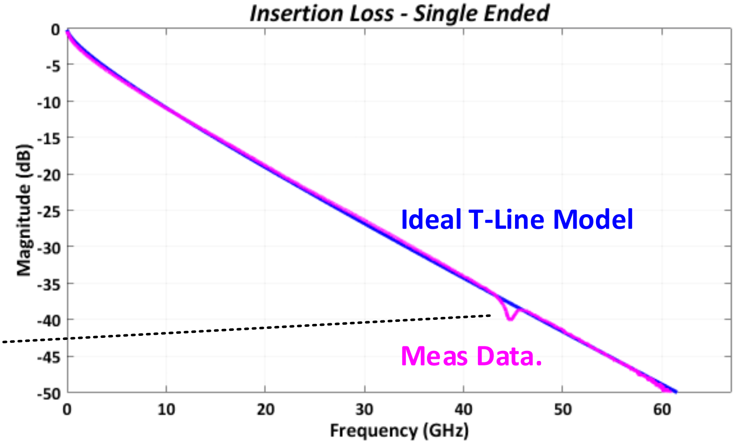
- Higher order TE modes excited by introducing center conductor step offset in simulation, observing notch response, and correlated to eigenmode solver
  - Step change in momentum
- Preserving modal BW requires shrinking the total conductor perimeter path length = higher loss

Fiber Weave Resonance – In Band Inconvenience



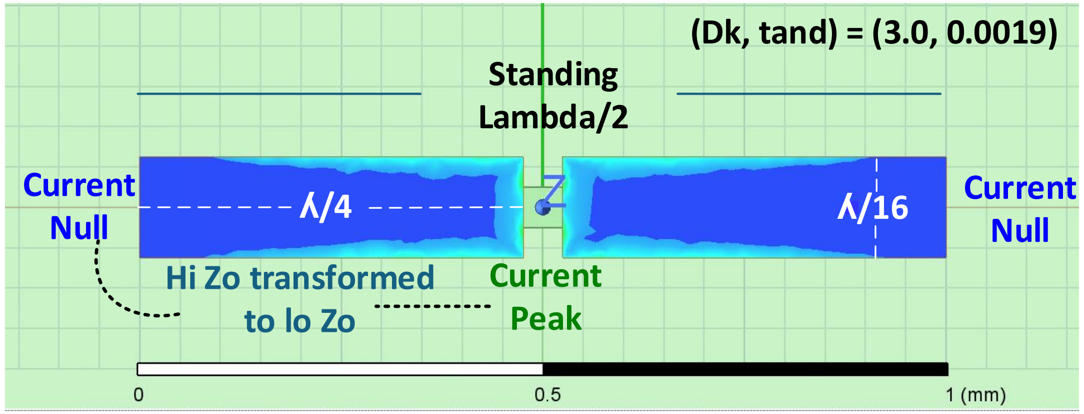


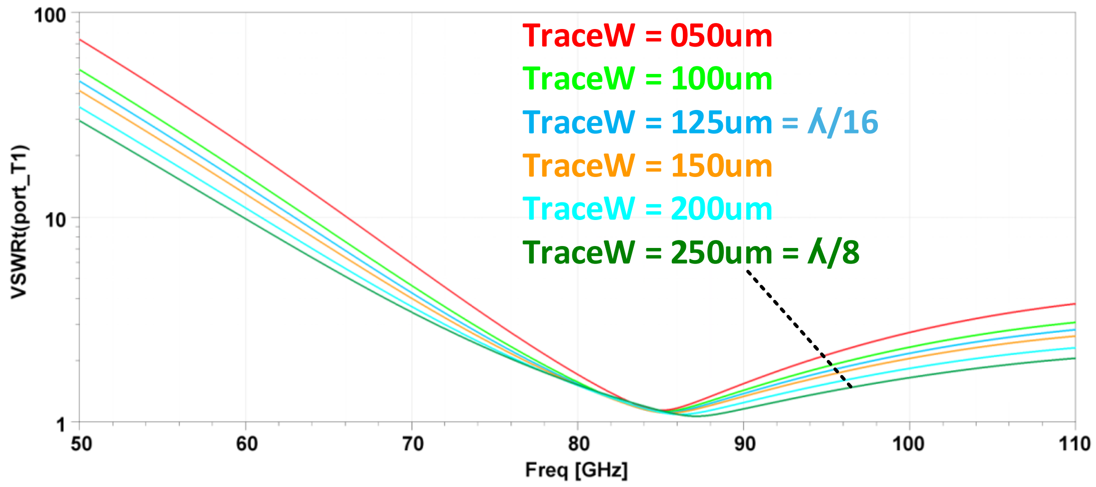


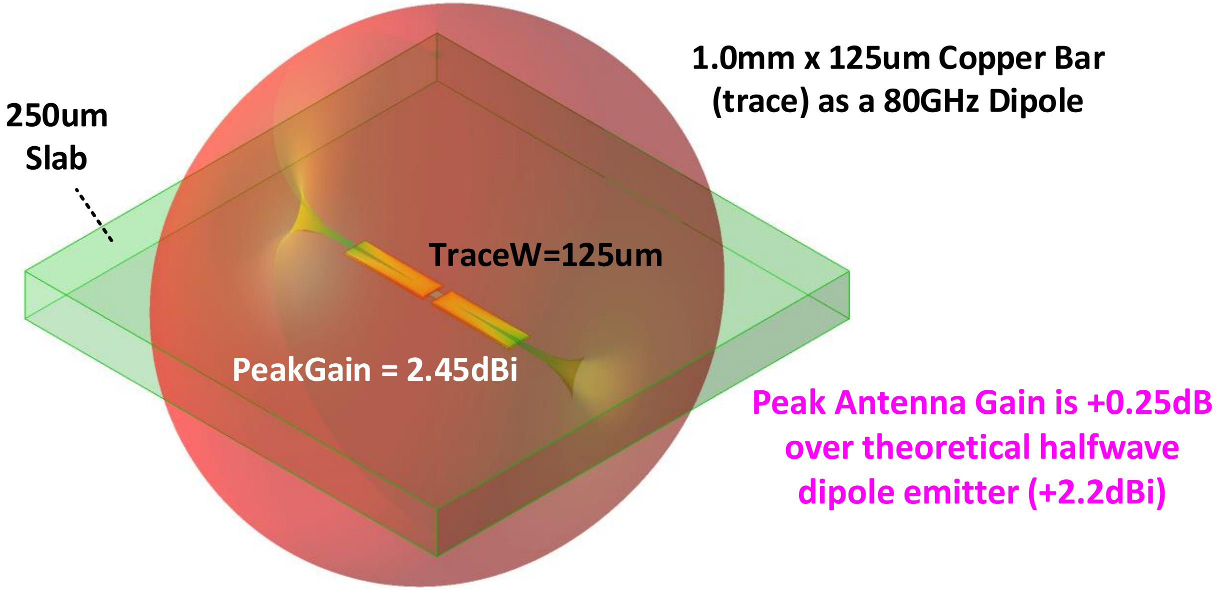


Apply Radiation Physics to PCB Interconnect Scale Features

PCB Striplines as Dipoles

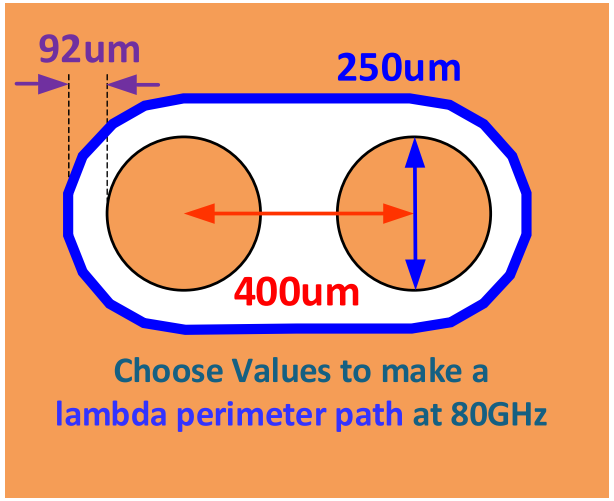






- Ask PCB scale interconnect radiation Questions.
  - When you act like a halfwave dipole…are you any good at it?
- By 250um trace width, standing half wave resonators have a VSWR=2 BW of 35% at ~2.5dBi antenna gain!
  - mm-Wave antenna designers are jealous

PCB Antipads as Slot Emitters

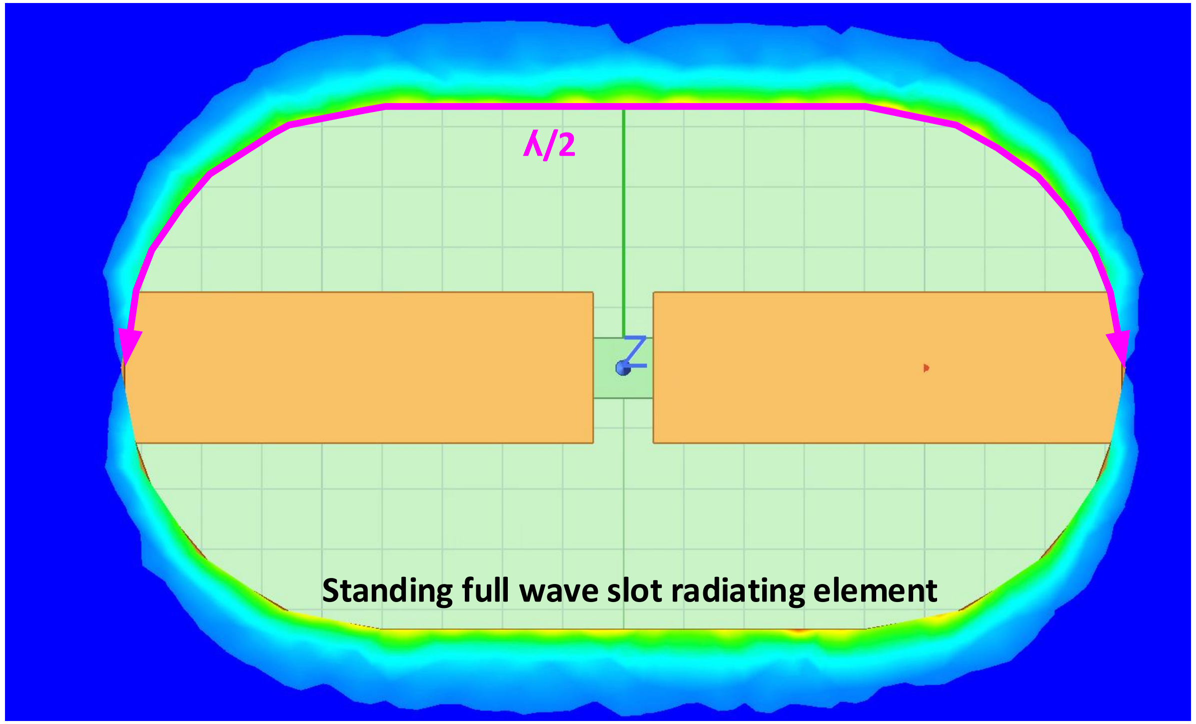


Choose Values to make a lambda perimeter path at 80GHz

	Technology		
Feature	PTH	HDI	SLP
P (mm)	0.600	0.400	0.250
D (mm)	0.400	0.300	0.175
G (mm)	0.100	0.100	0.075
er	3.000	3.000	3.000
Path (mm)	3.085	2.371	1.521
Fr	56.15	73.06	113.87

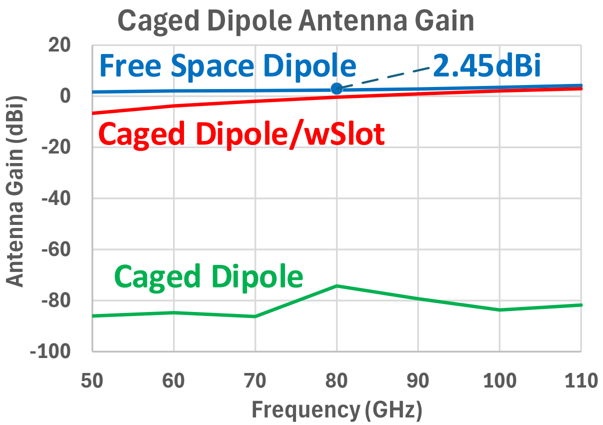
PTH = Single laminate PCB Technology
HDI = High density interconnect (stack laser drills)
SLP = Substrate like PCB

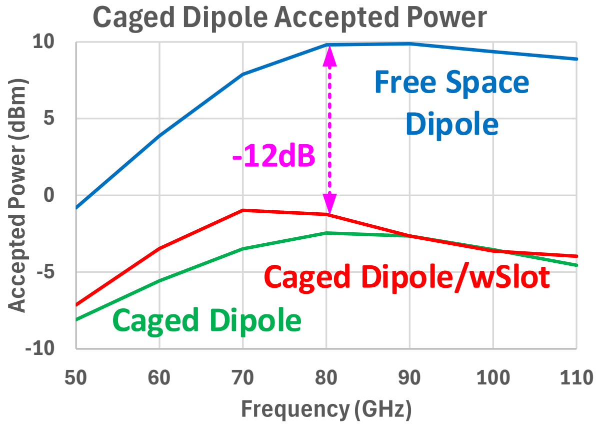
Excite the slot emitter by coupling to it from our 80GHz PCB Dipole Wires

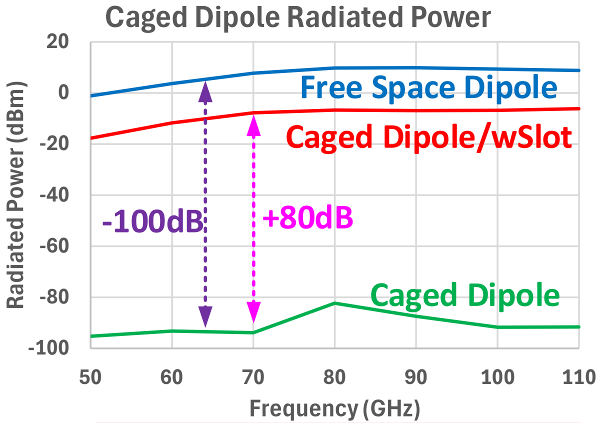


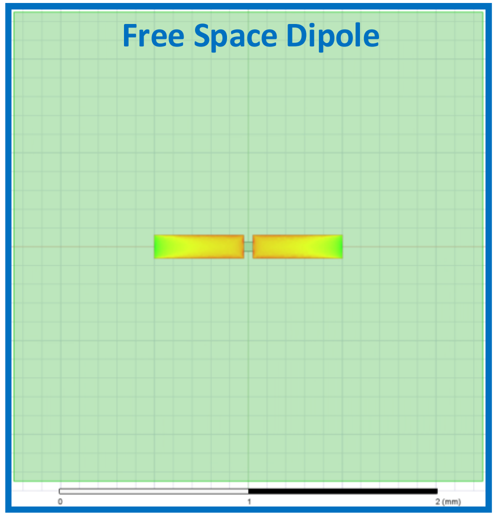
Halfwave resonance achieved between ‘vertices’ of the racetrack antipad

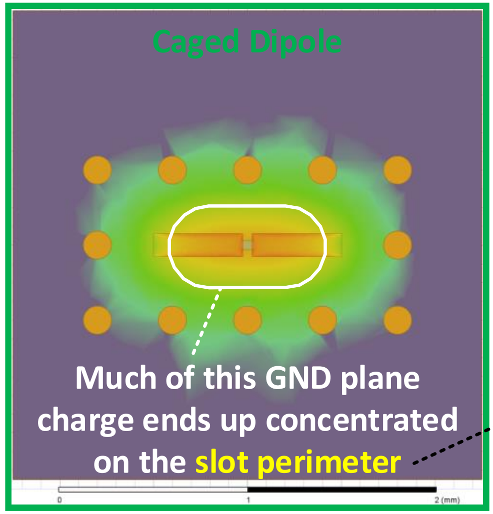
PCB Antipads as Slot Emitters


Caged Dipole Antenna Gain

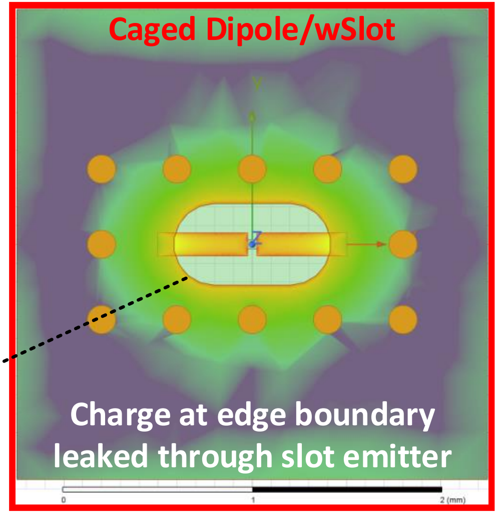

Caged Dipole Accepted Power


Caged Dipole Radiated Power


Free Space Dipole


Caged Dipole

Much of this GND plane charge ends up concentrated on the slot perimeter


Caged Dipole/wSlot

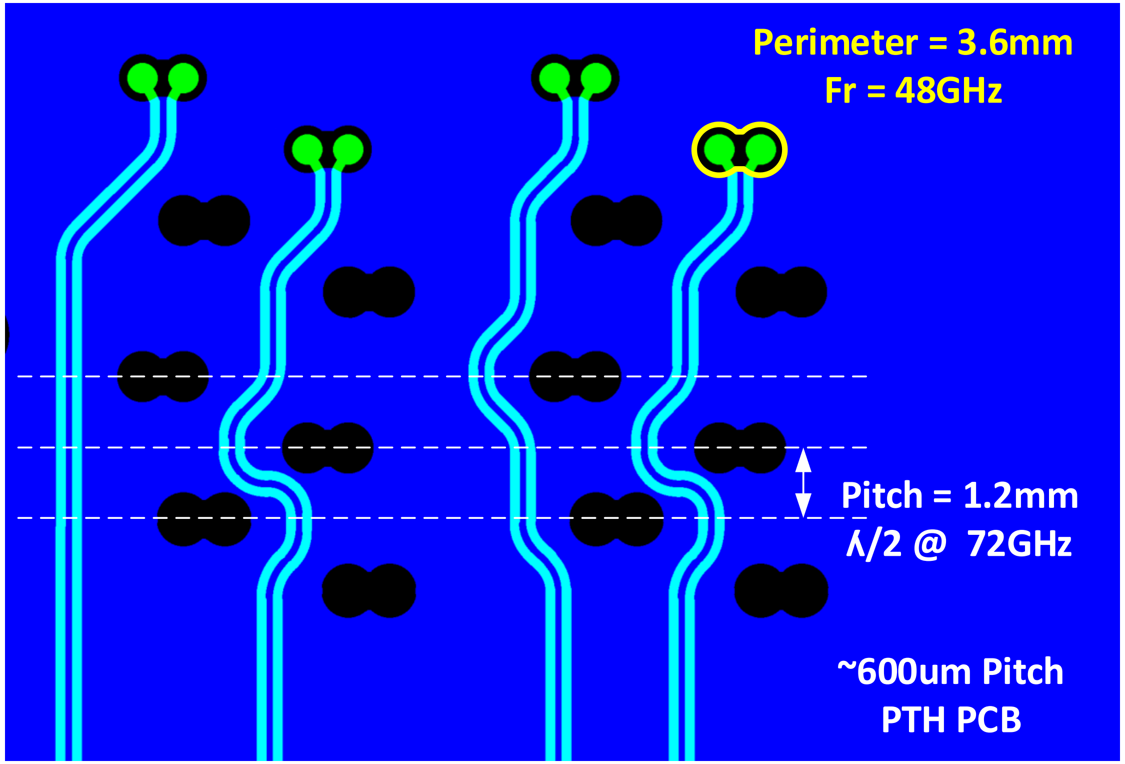
Charge at edge boundary leaked through slot emitter

Caging the dipole emitter in PCB ground planes with 400um pitch GND vias reduces radiated power by ~100dB

Adding a single lambda perimeter slot allows 80dB/100dB radiation suppression to leak through.

Still have not impedance matched the slot. The answer can get even worse.

PCB Antipads as Slot Emitters

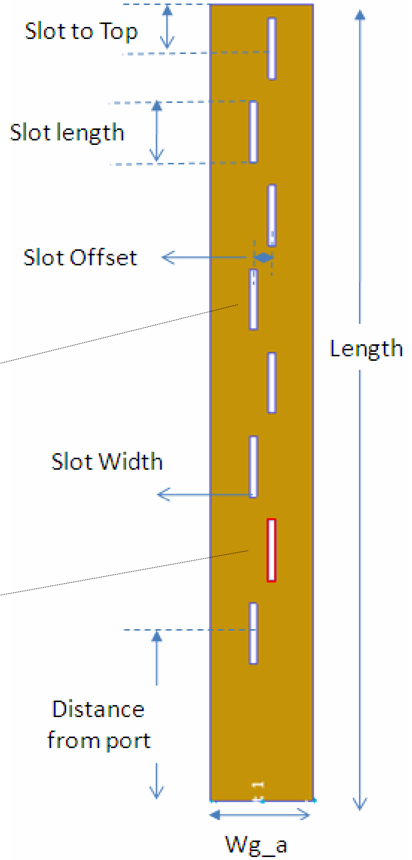


Perimeter = 3.6mm
Fr = 48GHz

Pitch = 1.2mm
λ/2 @ 72GHz

~600um Pitch
PTH PCB

- Constructive interference from multiple emitters is possible inside PCB component footprints when the signaling pitch exceeds a half wavelength.

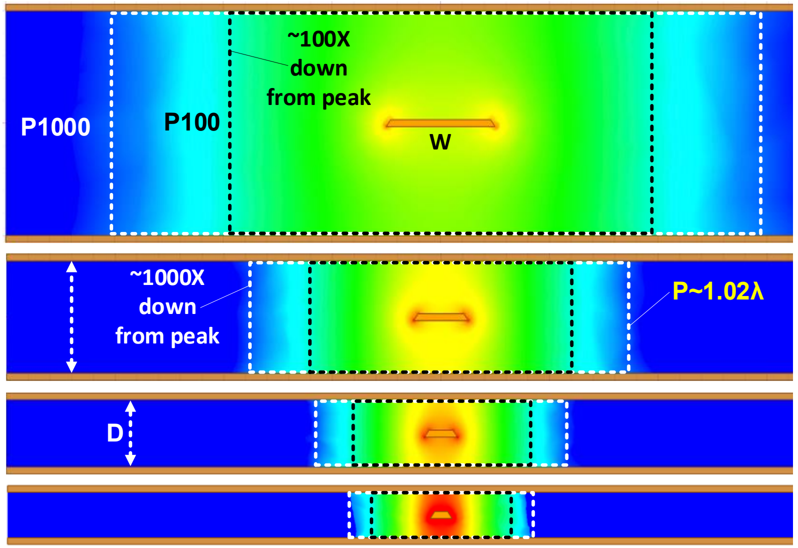


Slotted Array Antenna with λ/2 element spacing

Efficient resonant slot apertures

- Principle is exploited for high gain directional emitters in leaky waveguide arrays.

Modal Perimeter



Feature	Description
D (mm)	Ground plane to ground plane spacing
W (mm)	Stripline trace width for 50 Ohms SE
P100 (mm)	Modal perimeter path length in mm at ~40dB down from peak E-field magnitude.
P1000 (mm)	Modal perimeter path length in mm at ~60dB down from peak E-field magnitude.
P100 (80GHz Λ)	P100 in fractions of 80GHz wavelength
P1000 (80GHz Λ)	P100 in fractions of 80GHz wavelength

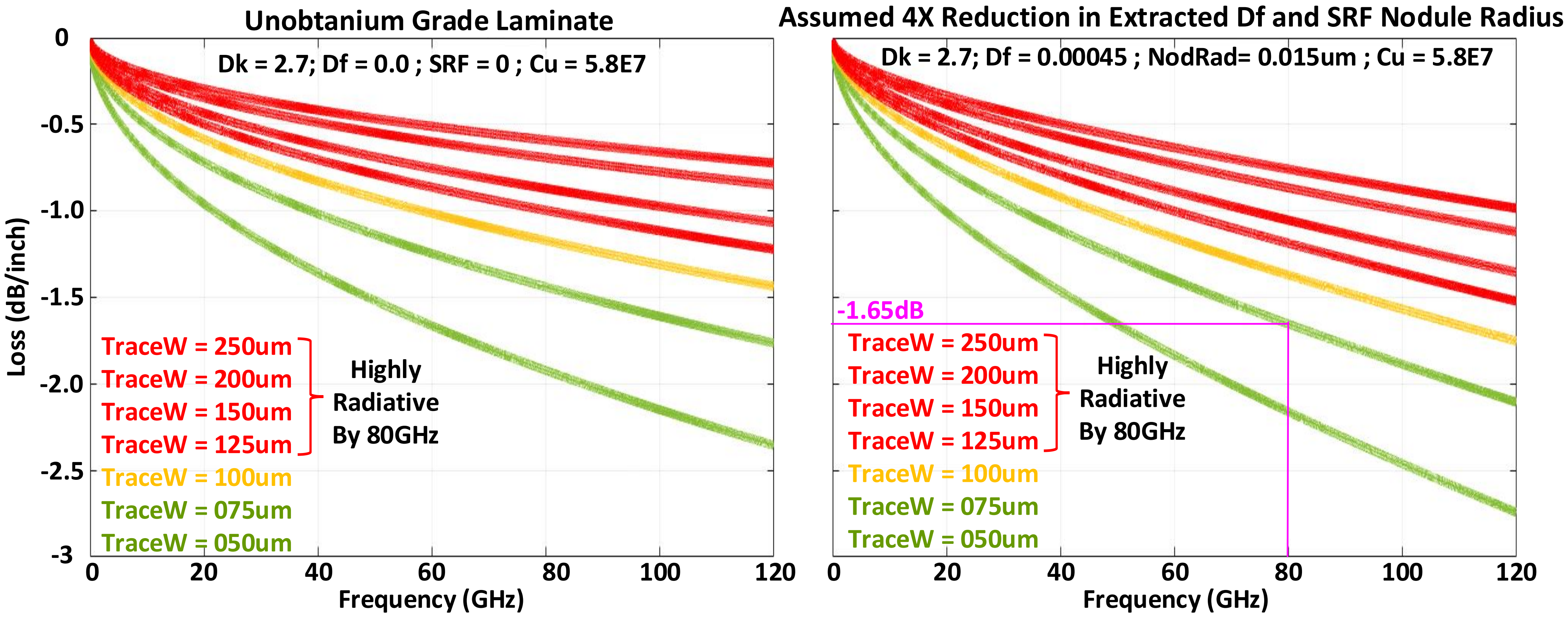
D (mm)	W (mm)	P100 (mm)	P1000 (mm)	P100 (53GHz Λ)	P1000 (53GHz Λ)	P100 (80GHz Λ)	P1000 (80GHz Λ)	P100 (106GHz Λ)	P1000 (106GHz Λ)
0.500	0.250	2.94	3.99	0.89	1.20	1.34	1.81	1.77	2.40
0.250	0.125	1.71	2.25	0.51	0.68	0.78	1.02	1.03	1.35
0.150	0.075	1.11	1.45	0.33	0.44	0.50	0.66	0.67	0.87
0.100	0.050	0.98	1.05	0.30	0.33	0.45	0.48	0.59	0.63

Decreasing
Line Width

Perimeter Scale	Domain
> 1.00Λ	Highly Radiative
0.5Λ - 0.99Λ	Leaky
< 0.5Λ	Interconnect

- Concept of a stripline modal perimeter requires specifying a field strength rolloff due to open side boundaries.
- Measure the stripline modal perimeter for E-field 40dB and 60dB down from peak.
  - H-field produces identical results
- Terminating the stripline in Z-axis interconnect becomes highly radiative when the modal extents exceed wavelength dimensions.
  - Implies a forced space transformation

Bounding PCB Loss

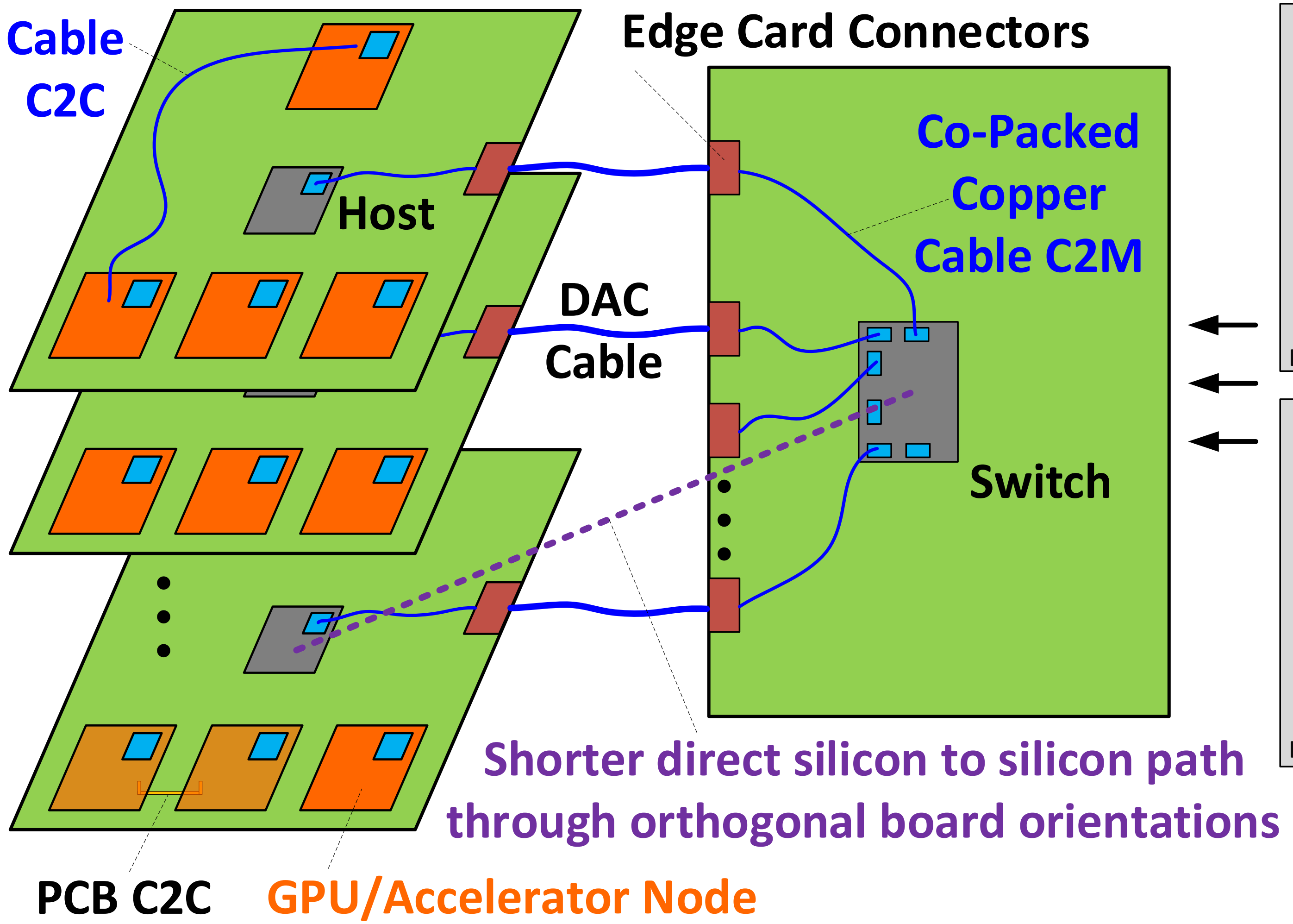

Unobtanium Grade Laminate

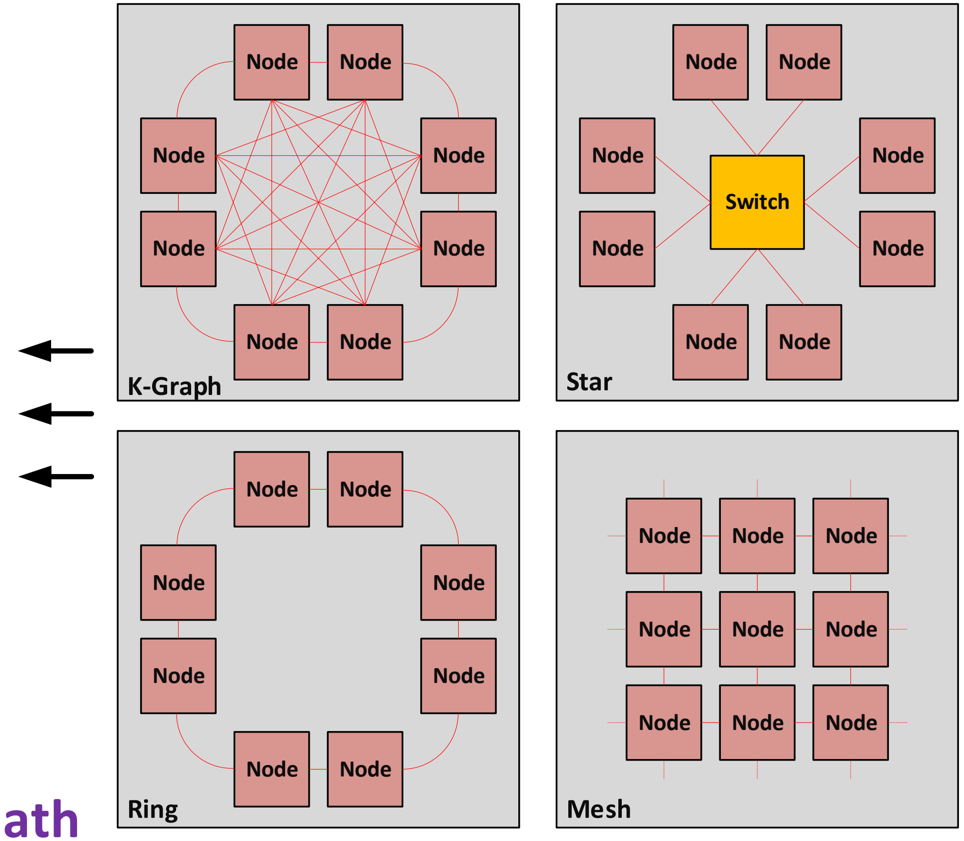
Assumed 4X Reduction in Extracted Df and SRF Nodule Radius

Maintaining 1.5dB/inch PCB stripline reach at Nyquist is likely not possible once radiation physical size constraints are overlaid on the planar interconnect.

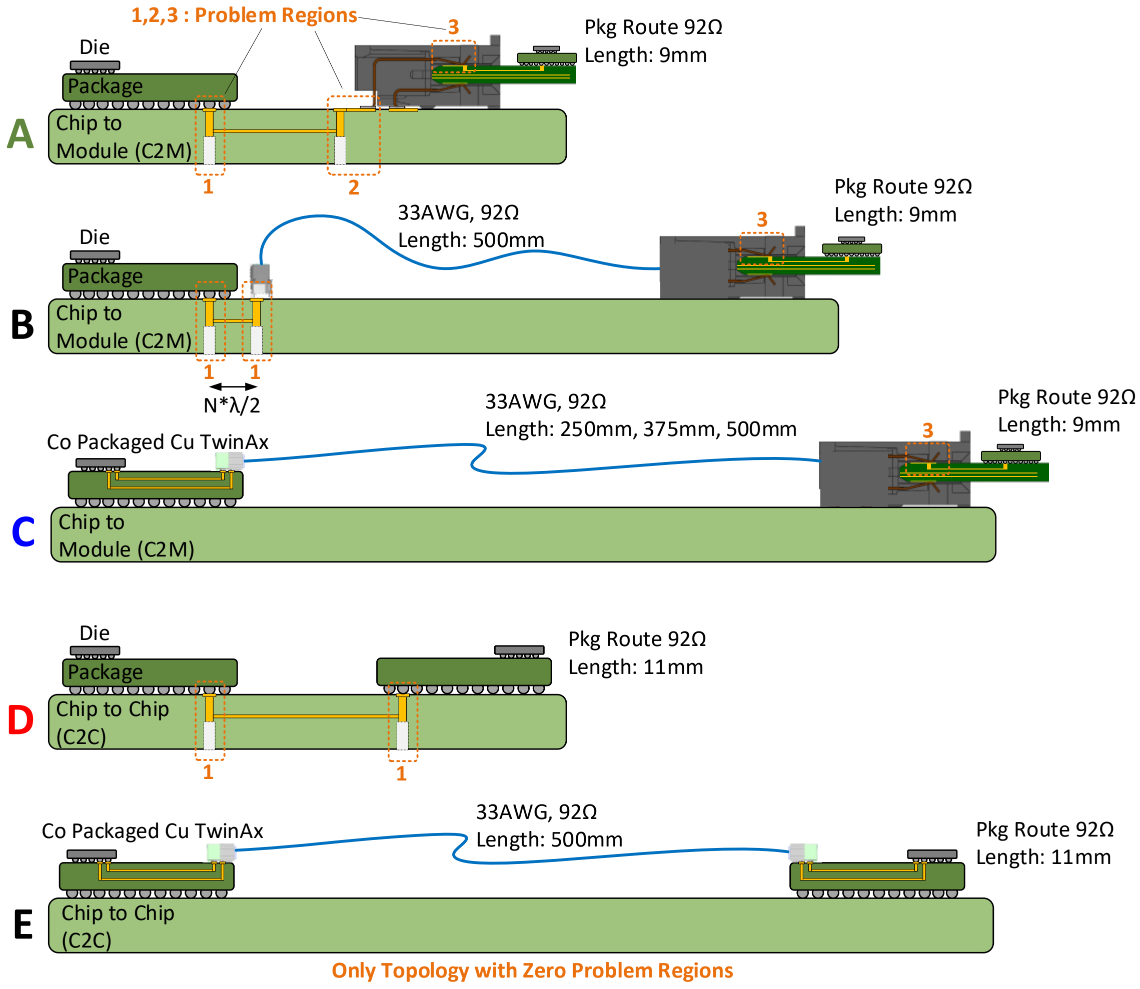
PCB Reach Bounded Topologies

Topology Considerations



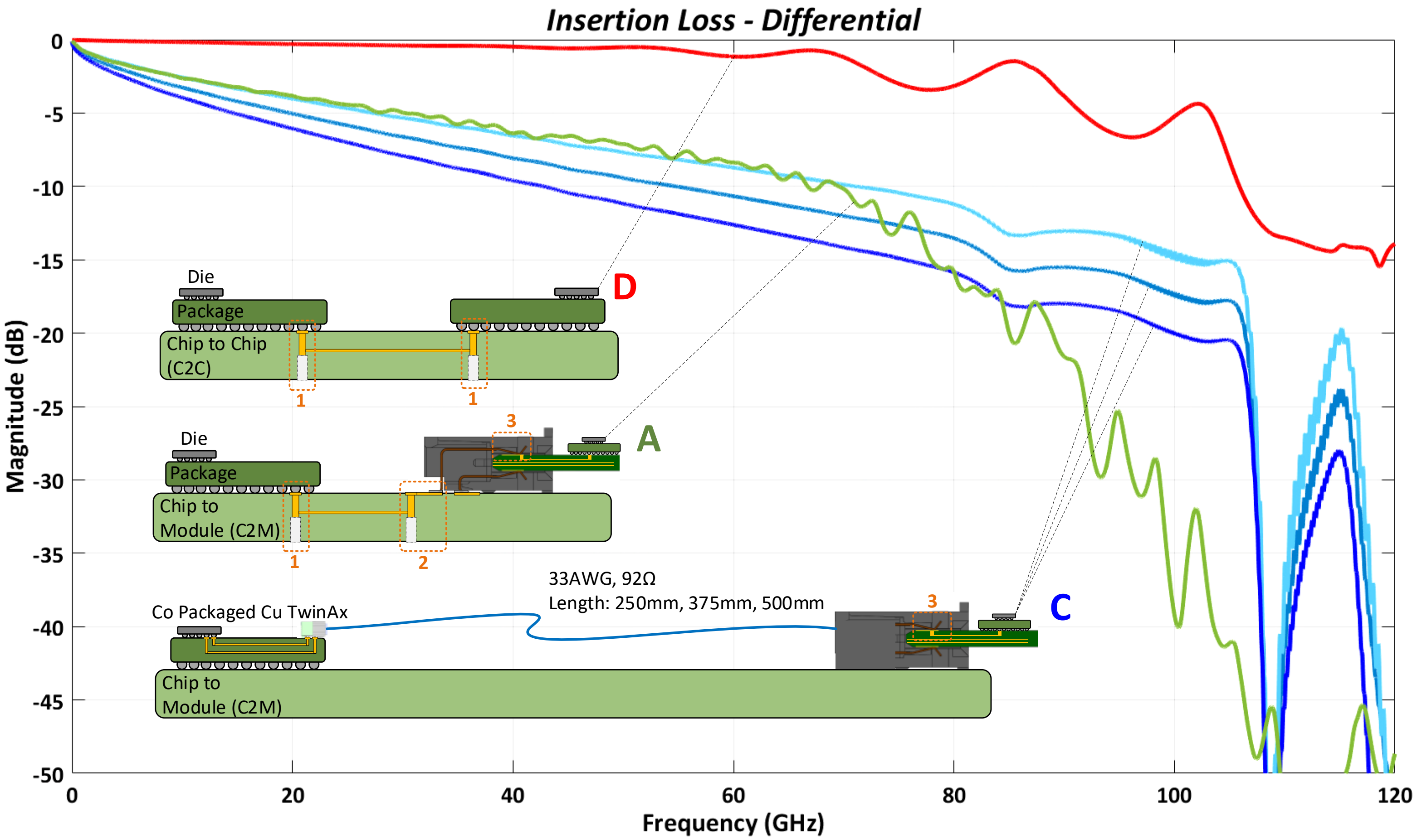


PCB Reach Bounded Channel Topologies



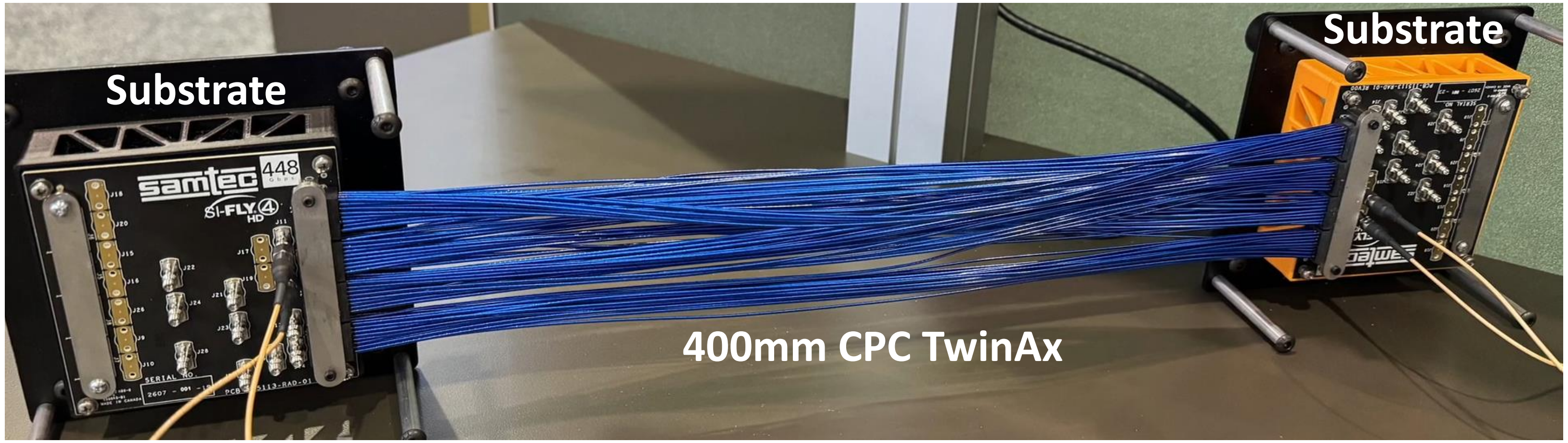
- (A) Reach limited PCB C2M Interface (traditional topology)
- (B) Reach extended Cabled C2M Interface
- (C) Reach Extended Co-packaged Cabled C2M Interface
- (D) Reach limited PCB C2C Interface
- (E) Reach extended Co-packed Cabled C2C Interface

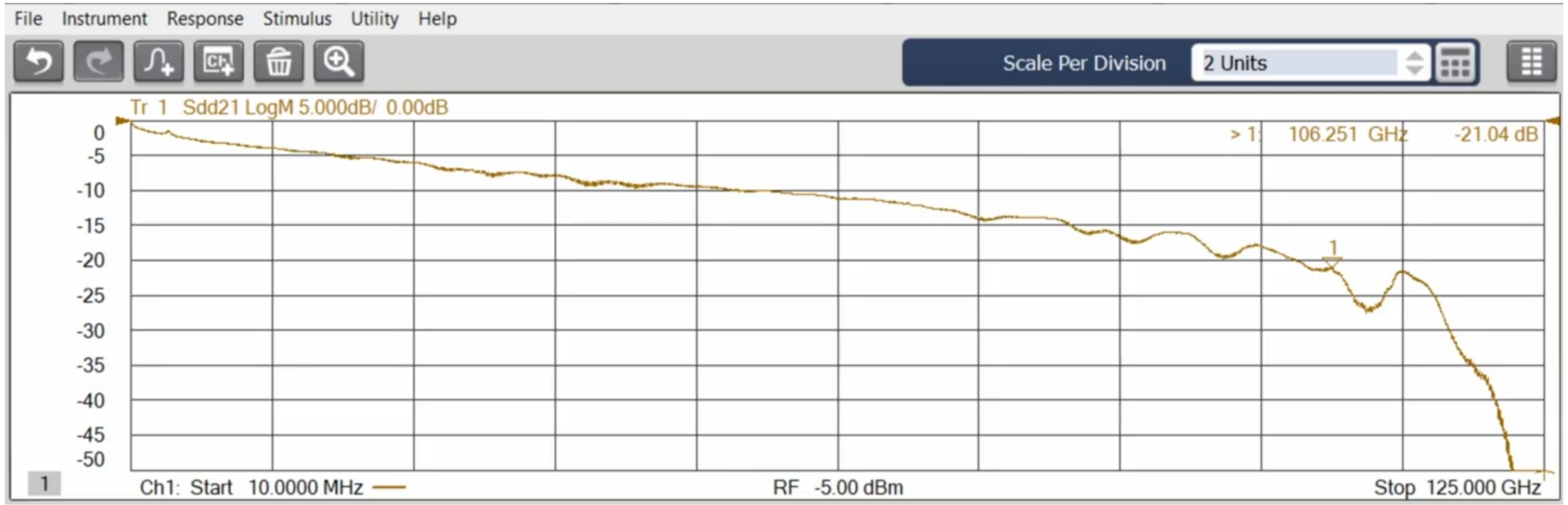
PCB Reach bounded Channel Topologies (Simulated)



- Pushing the rolloff frequency from the traditional **A-type** C2M topology over PCB requires eliminating reflective problem regions.
- These are the dominant mechanisms for enabling standing waves to be established on the interconnect that then can become efficient radiators in new signaling spectrums of interest.
- The only topology that can substantially mitigate all 3 problem regions is the Co-Packaged Twinax C2C link.

CPC C2C Links For Accelerator Meshes (Measured)





Scale Per Division 2 Units

Tr 1 Sdd21 LogM 5.000dB/ 0.00dB

> 1 106.251 GHz -21.04 dB

Ch1: Start 10.0000 MHz

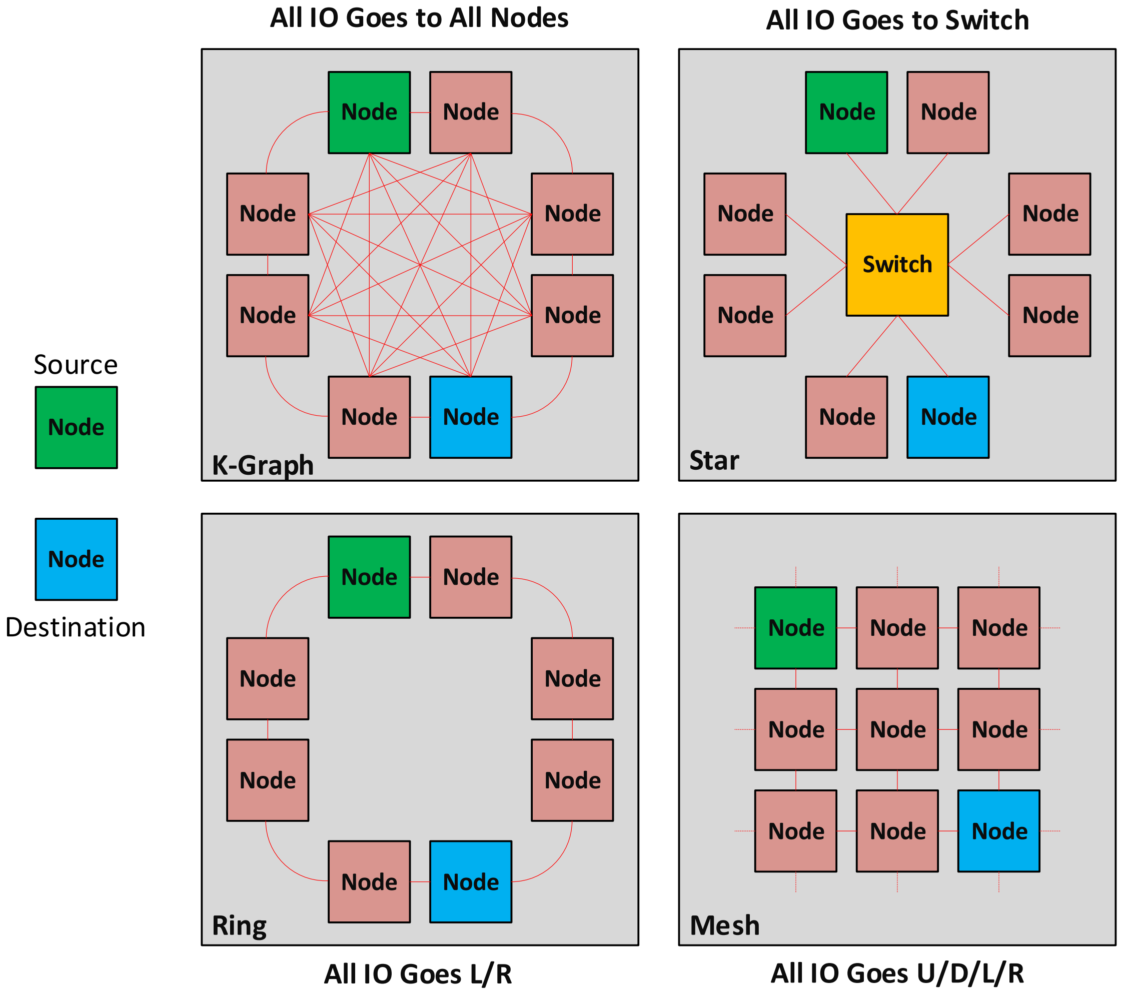
RF -5.00 dBm

Stop 125.000 GHz

- Co-Packed Cu Chip-2-Chip Link
- Targeted use case is accelerate node to accelerate node.
- 21dB Insertion Loss at 425Gbps PAM4 Nyquist (106.25GHz)
  - Includes ~2dB stripline wire loss from 1.00mm coax test point to CPC connector per side
  - 400mm of 92 Ohm 32 AWG Twinax

Compute Cluster BW Metrics

Network Hops in Parallel Computing



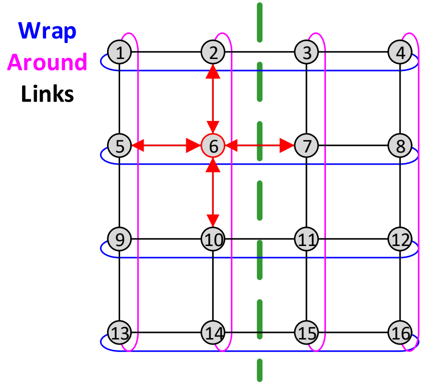
- Consider a parallel computing problem where each node in the cluster has local memory.
  - How does the cluster topology effect machine level BW and latency for node-to-node DMAs?
- Adjacent neighbor meshes work great for physics problems. Why?
  - Physics Problems -> Differential Equations
  - Continuous differentiable solutions
  - Great data spatial locality
  - If you need some data, chances are your neighbor has it.
- Ai compute is NOT a physics problem.
  - Poor data spatial locality.
  - If you need some data, chances are about equal anyone has it.
- What cluster level topologies are valuable?
- What cluster level BW metrics can we use?

Compute Cluster Bisection Bandwidth

$N^{1/2}$ scaling

(a)

Every Node talks NSEW

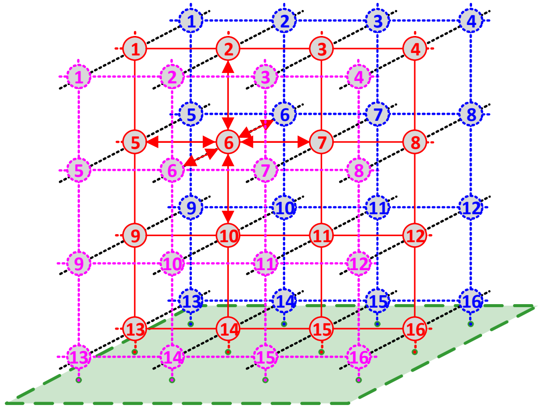


2D Torus Bisection Cut Line

- N = M*M
- Each node talks to 4 nearest neighbors
- Wrap Around Edge Links -> 2X
- *Num_Bisection_Link = 2*(N)^1/2*

$N^{2/3}$ scaling

(b) Every Node talks Up/Dn - NSEW

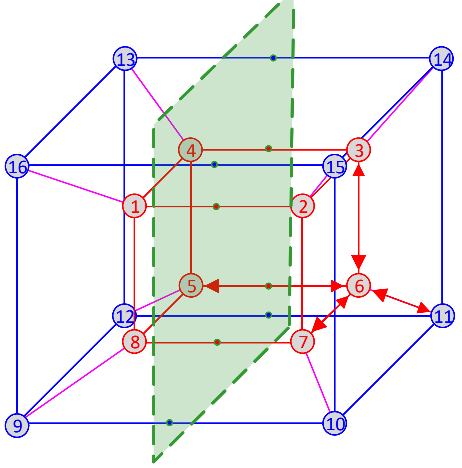


3D Torus Bisection Cut Plane

- N = M*M*M
- Each node talks to 6 nearest neighbors
- Wrap Around (not shown) Edge Links -> 2X
- *Num_Bisection_Link = 2*(N)^2/3*

N scaling

(c) Every Node talks to Log2(Nodes)

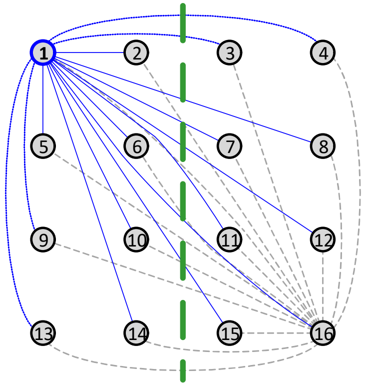


Hypercube Bisection Cut Plane

- N = 2^M
- Each node talks to M neighbors
- Each node sends 1 link across the bisection
- *Num_Bisection_Link = N/2*

N^2 scaling

(d) Every Node talks to Every Other Node

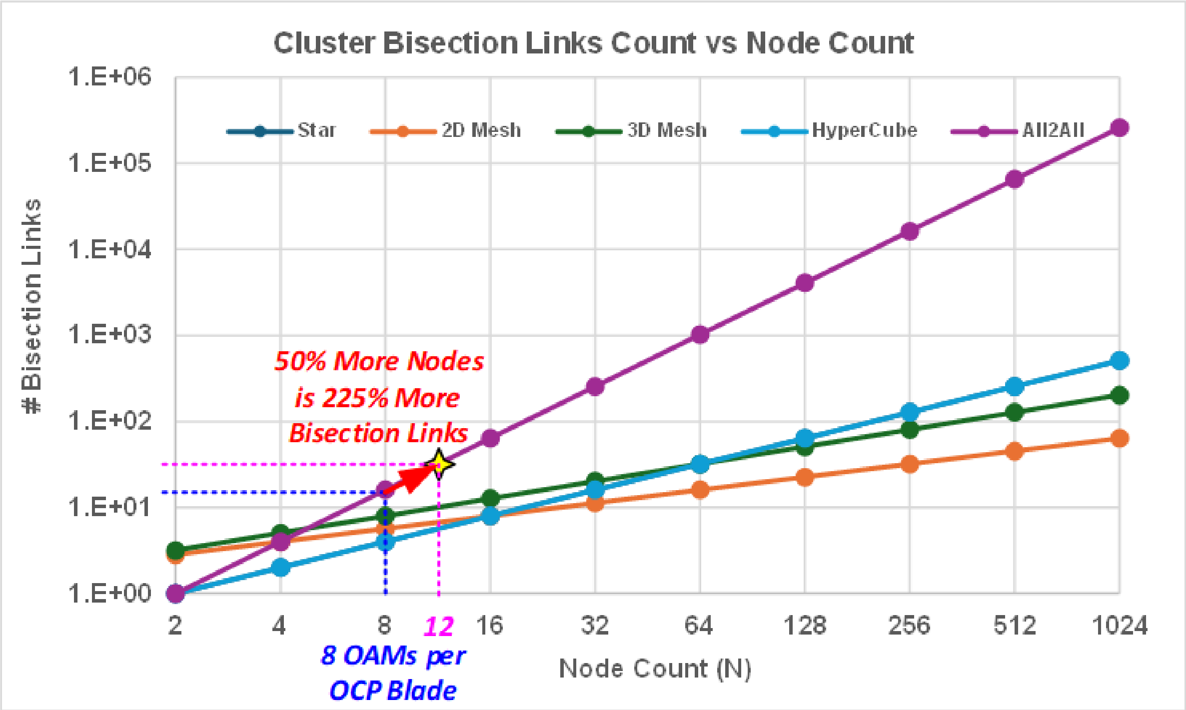


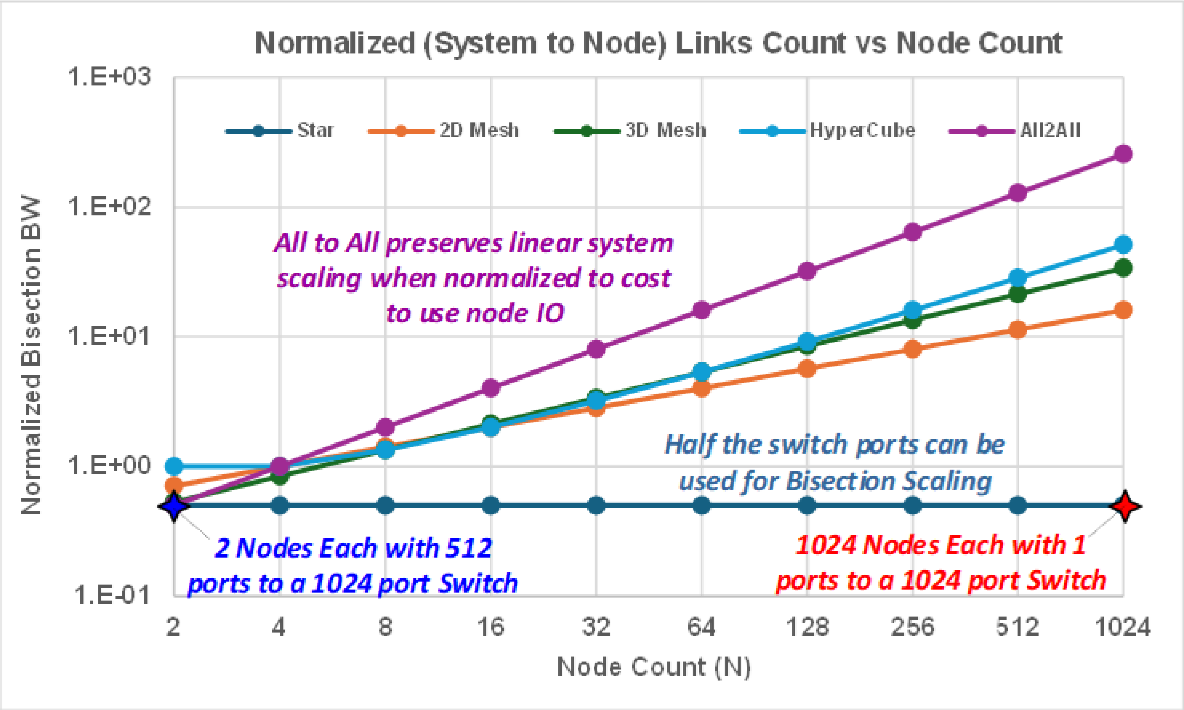
Flat All-to-All Bisection Cut Line
(2/16th connectivity shown)

- N = Integer
- Each node talks to N-1 nodes
- N/2 nodes from one half talk to N/2 from the other half N/2 times
- *Num_Bisection_Link = N*N/4*

- Bisection BW across the cluster is MIN bandwidth from half of the nodes to the other half.
- Different topologies have different Bisection BW scaling vs Node count.

Bisection BW Scaling vs Node Count





- Consider an accelerate blade with 8 nodes connected with 200G Class SerDes in an all-to-all network.
- Choice 1: Double link BW from 200G to 400G
- Choice 2: Increase mesh span by 1.414X

- Normalizing the cluster Bisection Links count to the # of available node IO provides a measure of IO efficiency in the cluster topology.

Conclusions

- The transition from 200G to 400G class serdes requires ~75GHz to 120GHz interconnect BW for the PAM4,6,8 modulations being considered by industry.
  - Nyquist ~ millimeter wave lengths
  - Pluggable interface mate mechanics become stub dominated from beam tips and edge card pads
  - Interconnect radiation mechanisms are mitigated with feature shrink
  - Feature shrink to transmission line cross sections increases insertion loss and limits physical reach

- Physical reach limits can be further mitigated through alternate HW topologies
  - Applies to both intra-blade (CPC C2C) and inter-blade (Cabled Backplane, Cabled Orthogonal Midplane) links

- What you do with the links in a parallel compute cluster matters as much as the link data rate.
  - An actionable hardware-centric definition of cluster bisection BW can be used as the system level overlay to extend Phy and link level work to machine level BW useful at the problem-class/algorithm level.
  - Allows SI Engineers to participate in Phy to Channel to Cluster Co-Design

Useful References

- Beyond 200G: Brick Walls of 400G links per Lane
  - Paper: Beyond 200G: Brick Walls of 400G Links Per Lane
  - Slides: Samtec DC25 Beyond 200G: Brick Walls of 400G Links Per Lane
- ManyPoint Networks
  - Paper: Cluster Topology and PHY Co-Design at 448 Gbps | Signal Integrity Journal
- Precision mm-Wave Test Fixture for 400Gbps Serial PHY Characterization
  - Webinar: Precision mm Wave Test Fixtures for 400Gbps Serial PHY Characterization

SUDDEN SERVICE®