56G/112G Link Foundations
Standards, Link Budgets & Models

Ganesh Balamurugan, Ajay Balankutty,
Chun-Ming Hsu

Intel

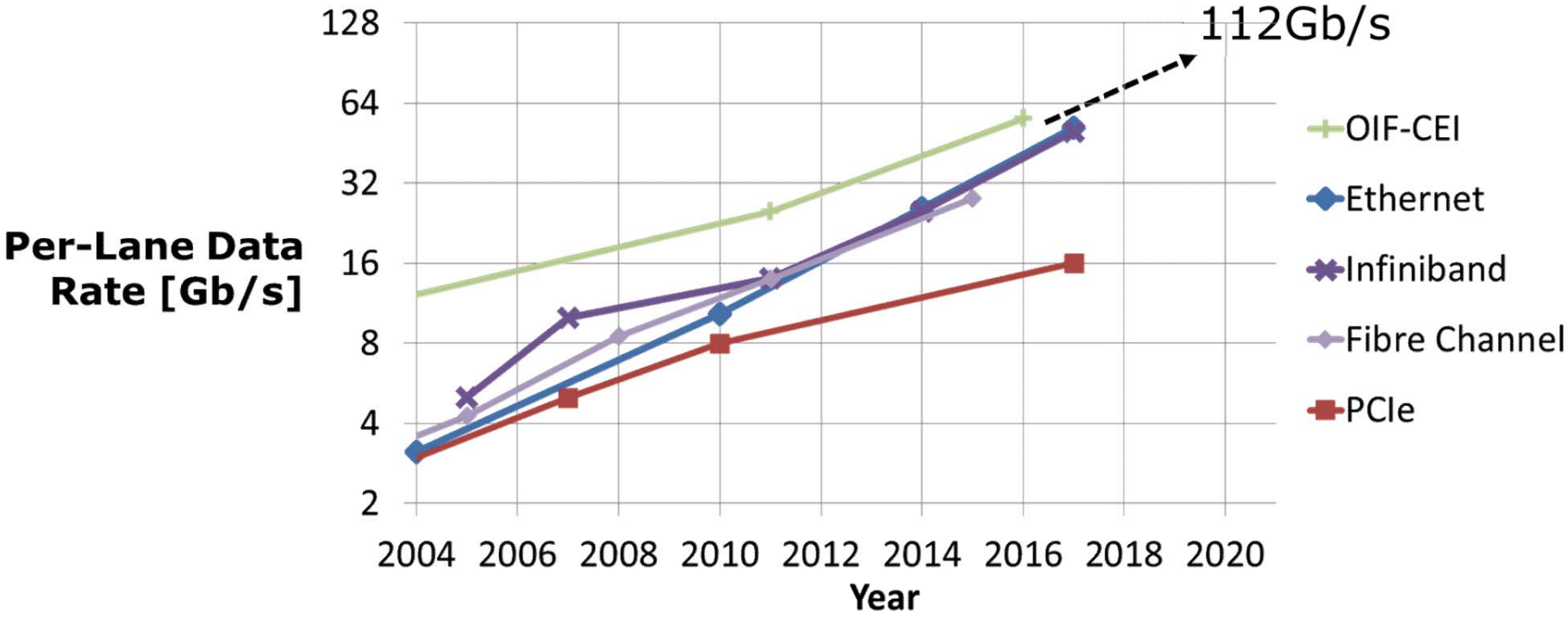
Outline

- Introduction
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

Outline

- Introduction
  - Wireline trends and drivers for scaling
  - Example 400G data center link
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

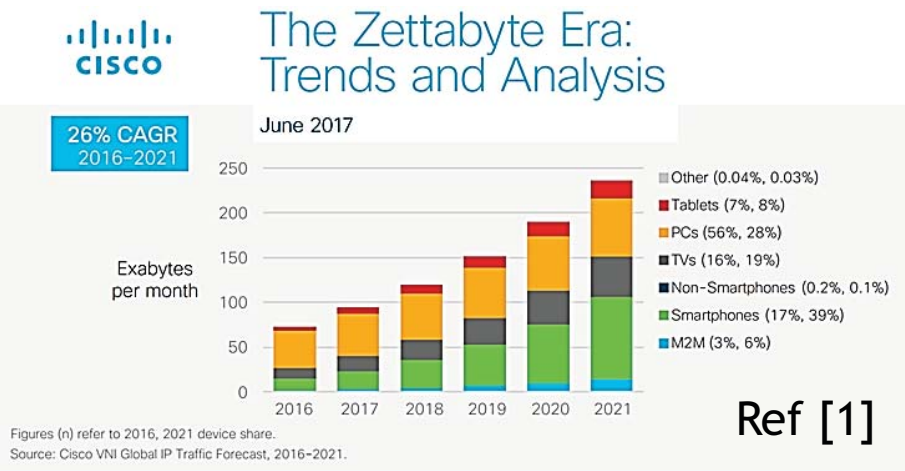
Wireline Data Rates (2004-2018)



- Wireline rates have doubled every 3-4 years
- 28Gb/s deployed, 56Gb/s deployment underway, 112Gb/s in development

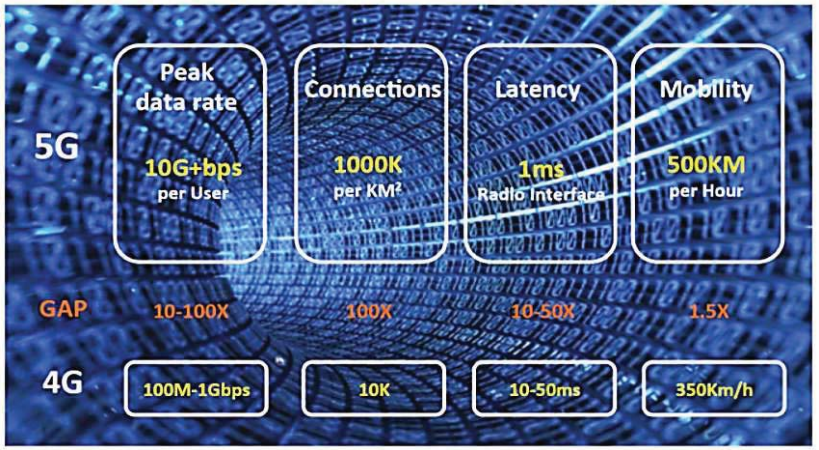
Drivers for Bandwidth Scaling

Internet Traffic Trends



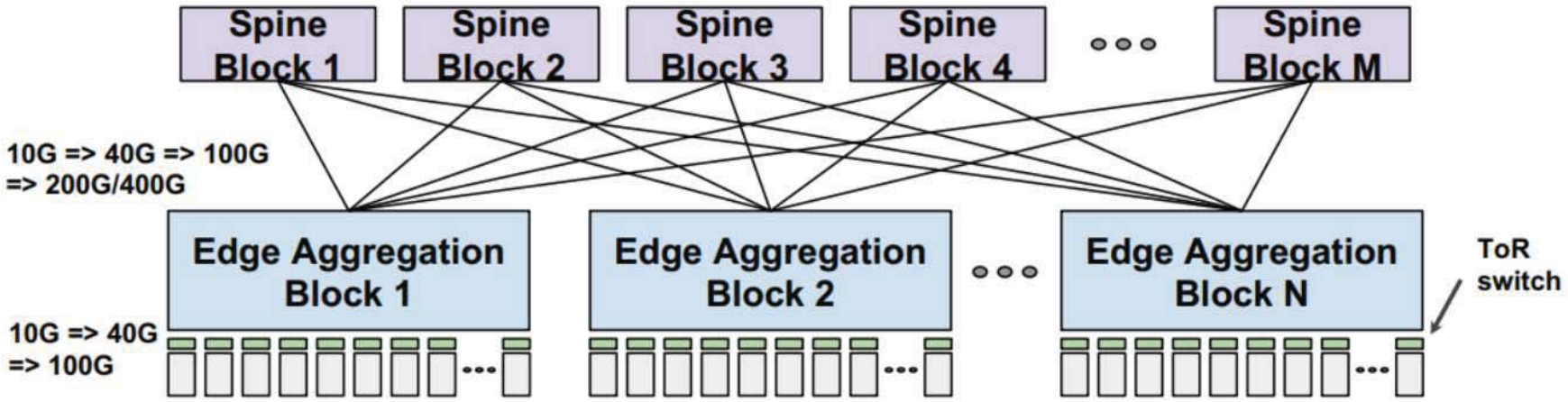
Ref [1]

4G-to-5G Transition



- Internet traffic and high-performance computing drive bandwidth scaling
- Further bandwidth leap with transition to 5G
- Data centers evolving to meet this demand

Data Center Trends



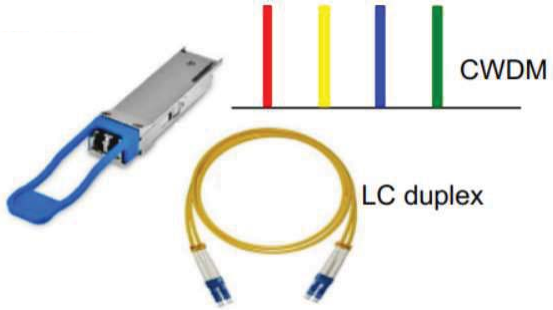
Ref: "A Decade of Clos Topologies and Centralized Control in Google's Data Center Network," Amin Vahdat, ONS 2015 Keynote

- Network bandwidth grew by 100X over a decade [2]
- Key hardware elements to sustain scaling:
  - Switch silicon
  - Electrical and optical interconnects

Interconnects in Data Center

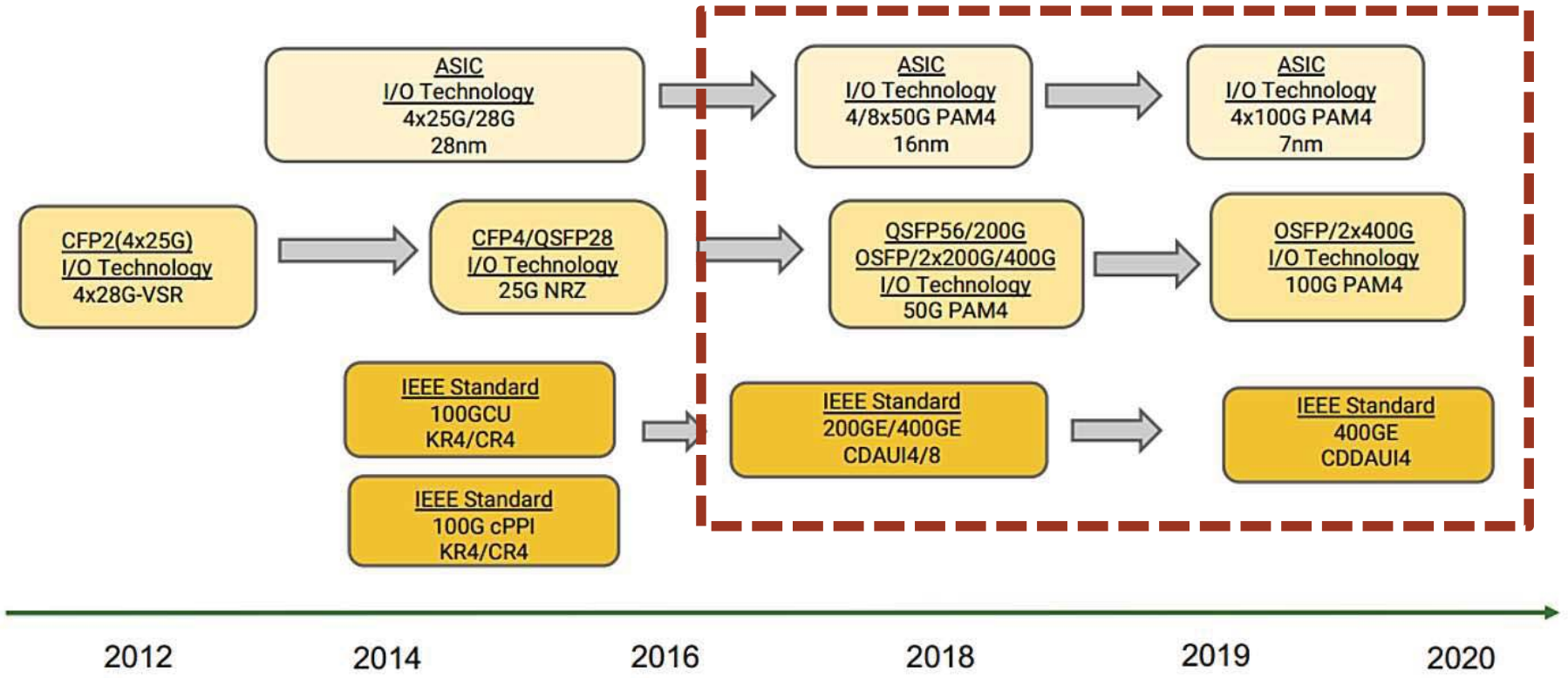
Reach for 100Gb/s 2m
Reach for 50Gb/s 3m

x12 parallel fibers



- Electrical within rack (<3m):
  - Direct attach copper (DAC)
  - Gauge varies with reach
- Optical outside rack:
  - For <100m: Parallel MMF (OM3, OM4), VCSEL-based
  - For 100m-2km: Duplex SMF, WDM, SiPh-based

I/O Evolution for Data Center Optics

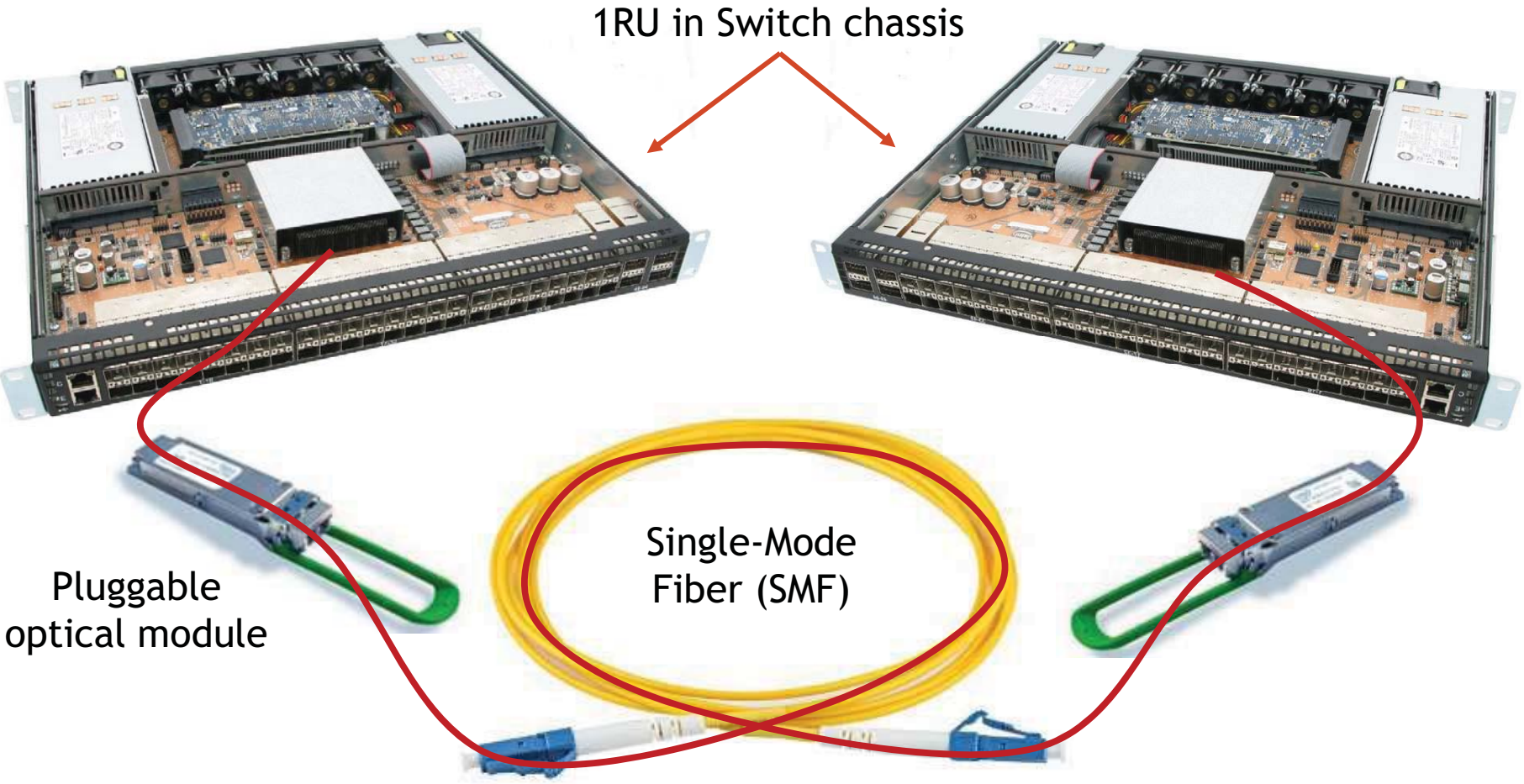


- I/O standards and ASICs evolving to meet Data Center (DC) bandwidth demand
- Focus of this talk: Links with 50G-100G per-lane rates

Outline

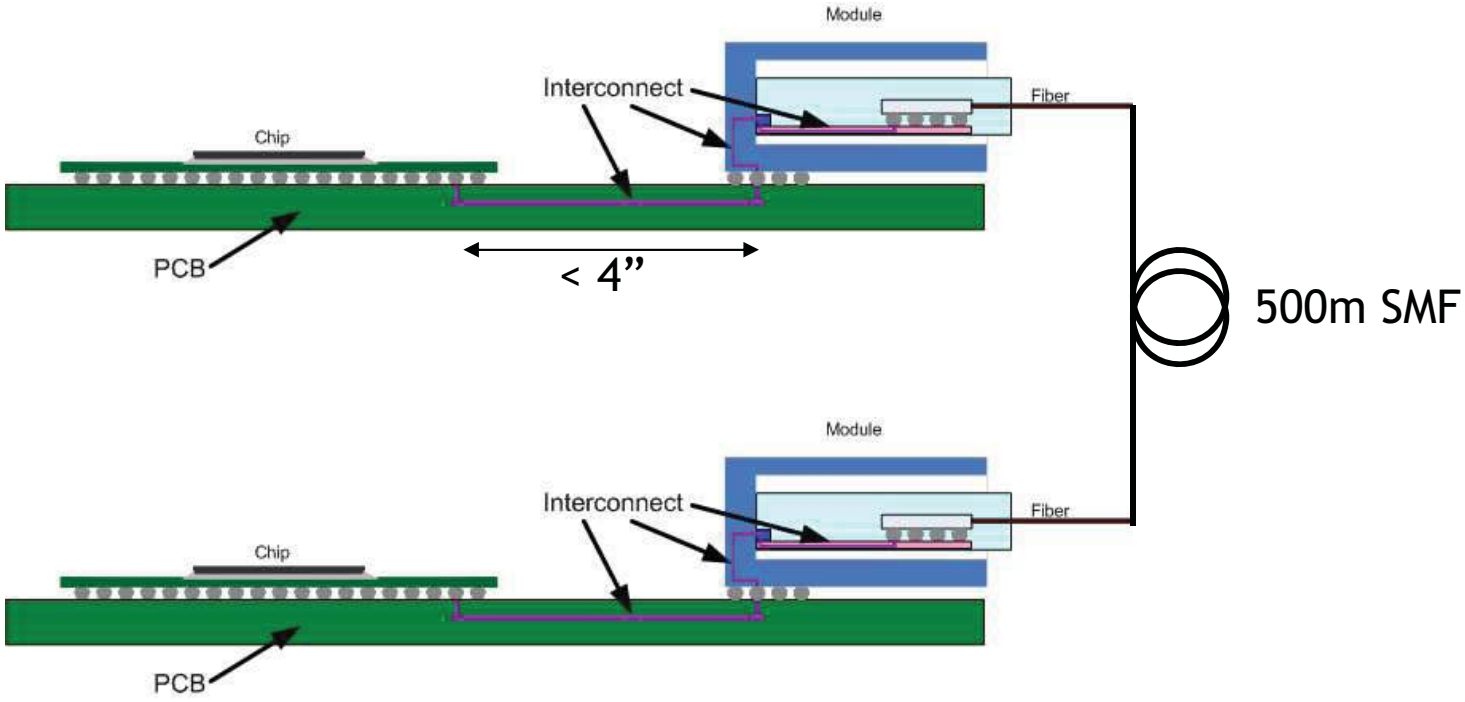
- Introduction
  - Wireline trends and drivers for scaling
  - Example 400G data center link
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

Example 400G DC Link - Physical View



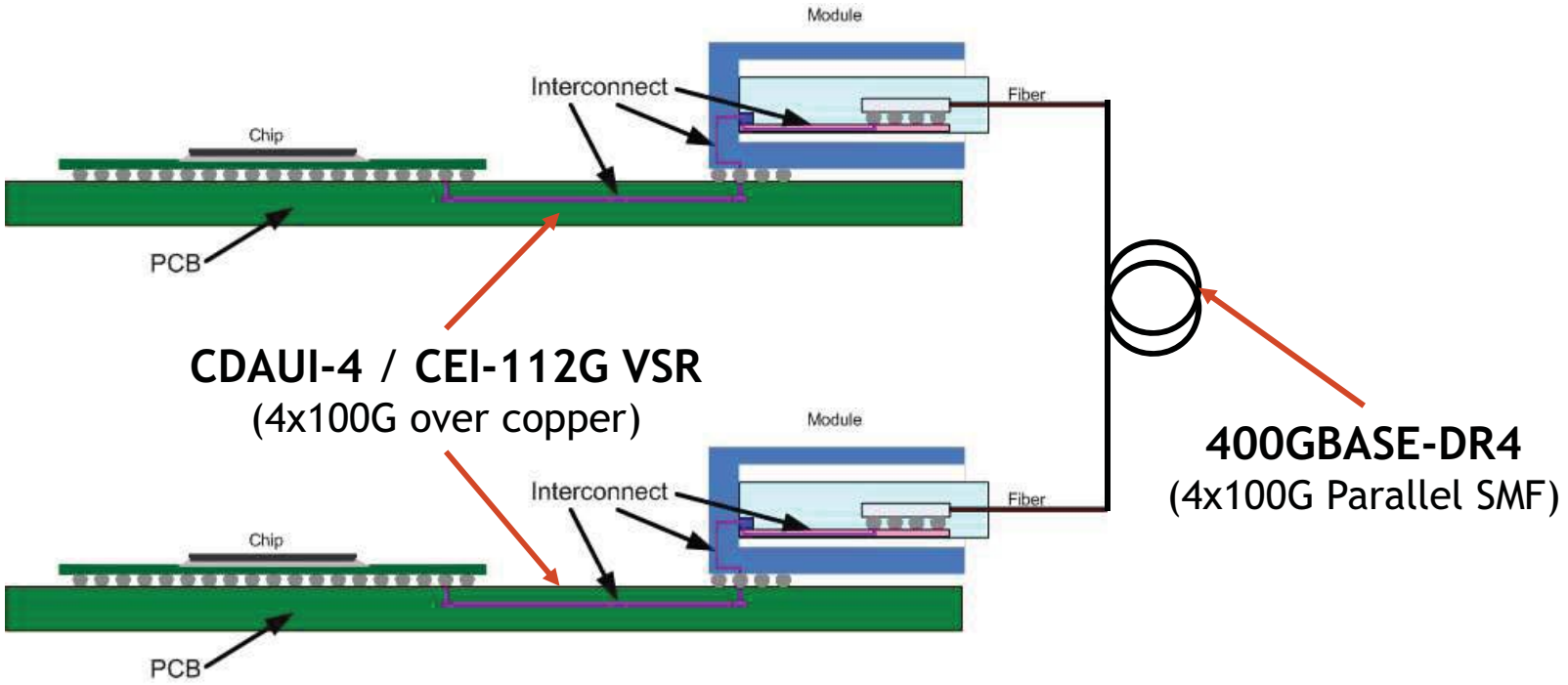
- 400G Inter-rack data center link with ~500m reach

Example 400G DC Link - Schematic View



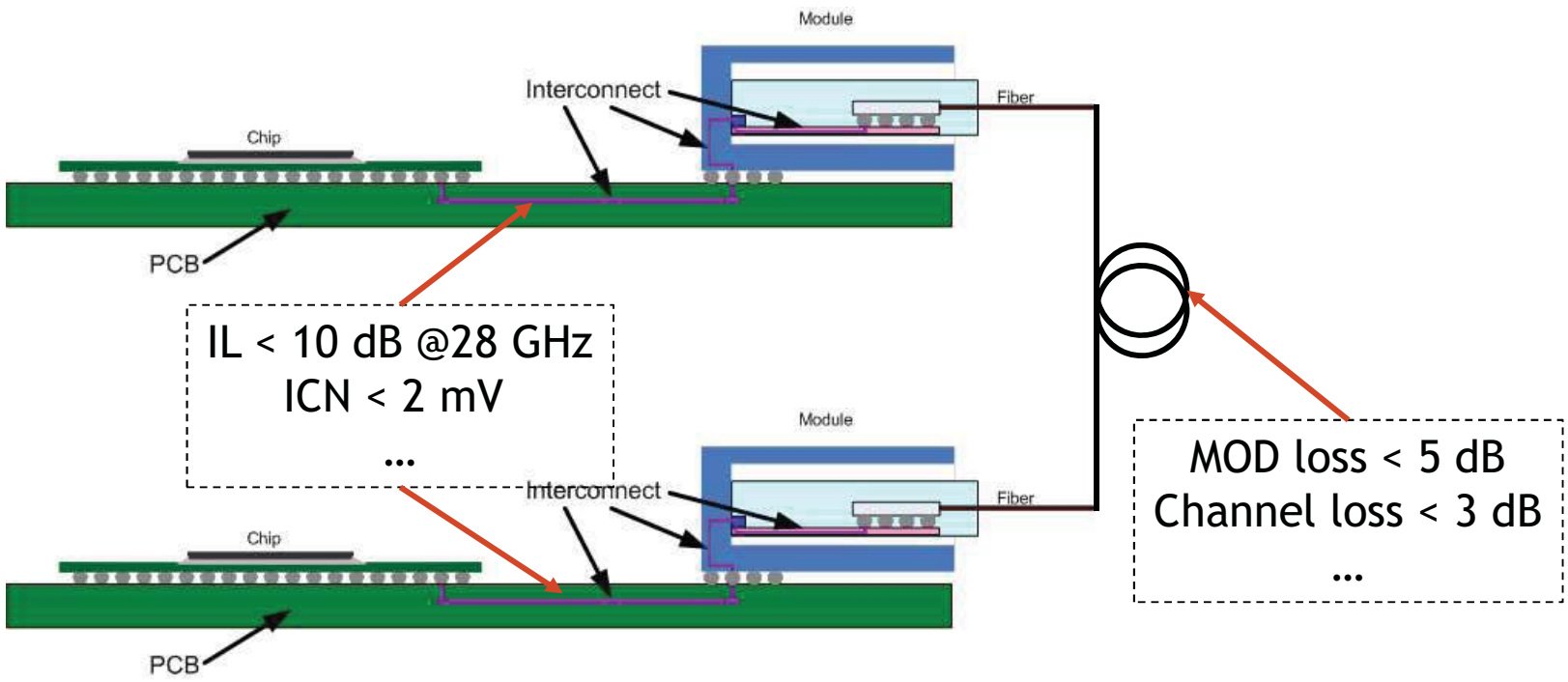
- Electrical link: Chip-to-Optical Module
- Optical link: Module-to-Module through SMF

Example 400G DC Link - Standards



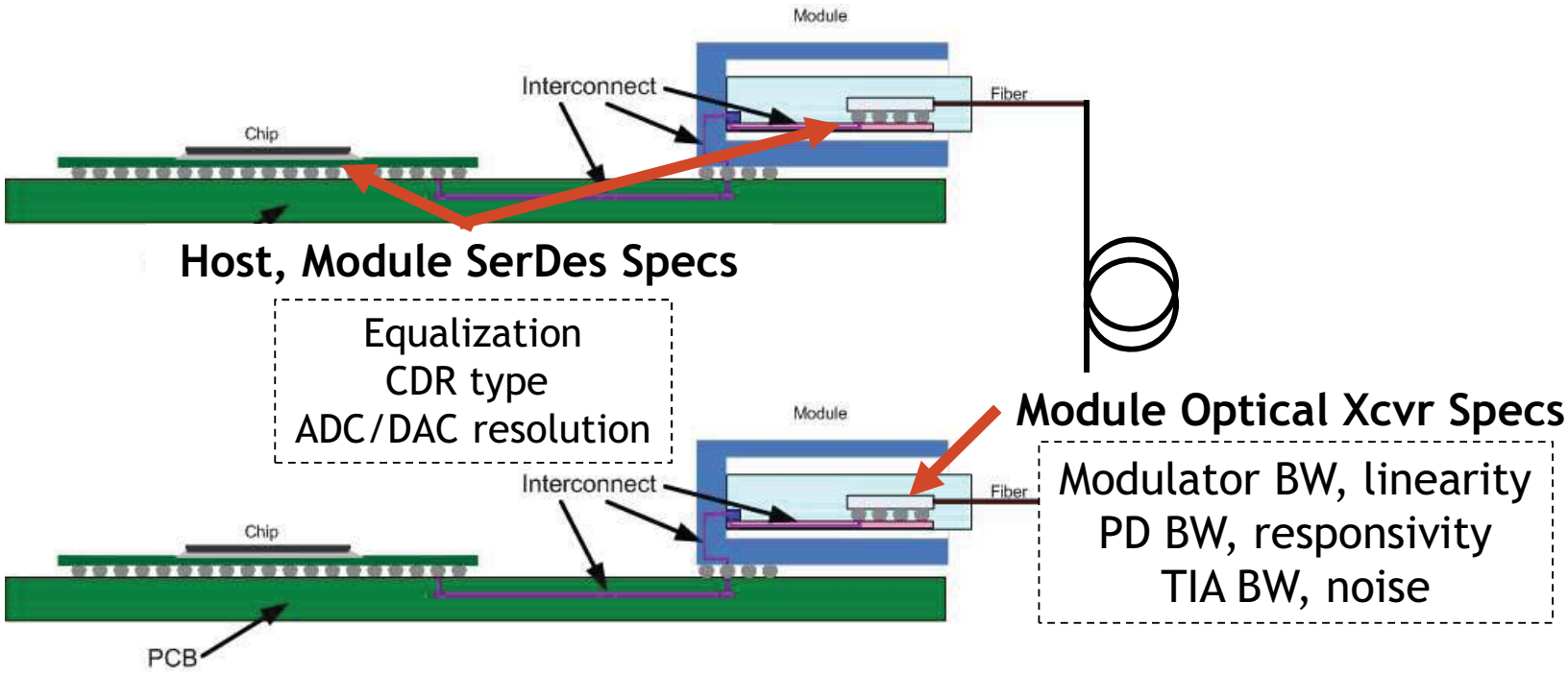
- Standards specify broad requirements for interoperability
  - E.g. Baud rate, modulation, target BER

Example 400G DC Link - Link Budgets



- Link budgets specify tolerable limits for impairments to meet Standards' specs
  - E.g. insertion loss, crosstalk, optical modulator penalty …

Example 400G DC Link – Link Models



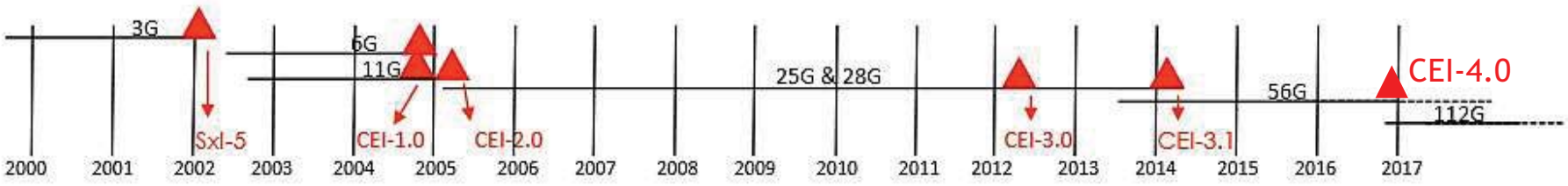
- Link models are used to:
  - Explore SerDes, optical transceiver architectures
  - Determine circuit- and device-level specifications
  - Validate end-end link performance

Outline

- Introduction
- 56G,112G Standards
  - Background
  - Electrical standards
  - Optical standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

Wireline Signaling Standards

- Criteria for standards development:
  - Broad market potential
  - Technical *and* economic feasibility
- Goal: Develop implementation agreements (IAs) to facilitate interoperability
- 2 Key standards bodies:
  - Optical Internetworking Forum (OIF) [3]
  - IEEE 802.3 (Ethernet) [4]



56G/112G Electrical & Optical Standards

Electrical
(On-package to cable/backplane)

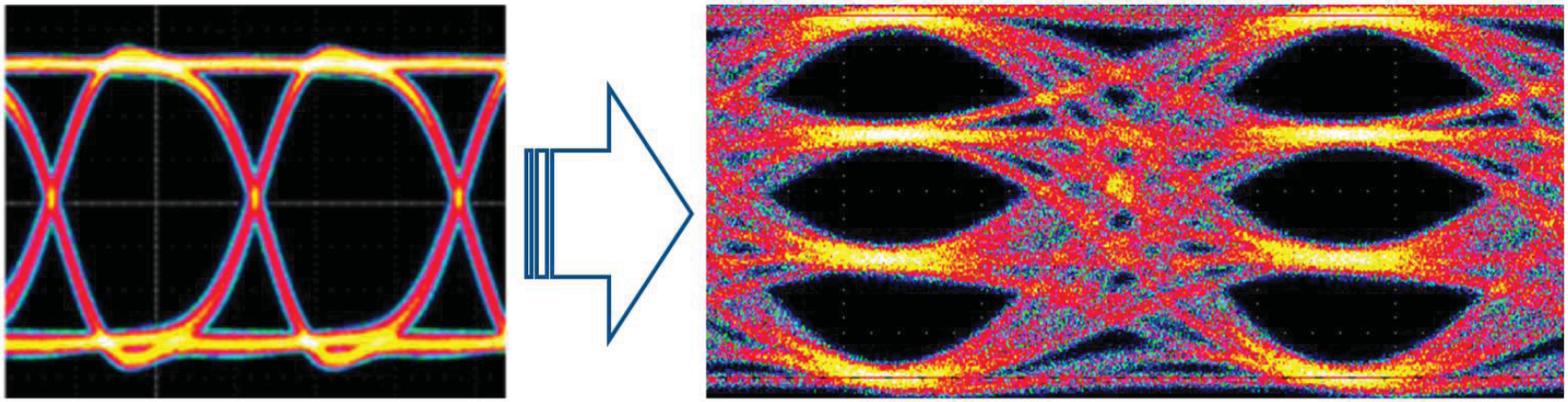
- OIF-CEI [5]
  - CEI-56G-XSR (on-package)
  - CEI-56G-LR (backplane)
  - …
- IEEE Ethernet (802.3) [6]
  - 200GAUI-4
  - 200GBASE-KR4
  - …

Optical
(100m-10km)

- 200G/400G Ethernet (802.3bs) [6]
  - 200GBASE-FR4/LR4 for 2/10km
  - 400GBASE-DR4 for 500m
  - …
- 50G/100G/200G Ethernet (802.3cd)
  - 200GBASE-SR4 for MMF
  - …

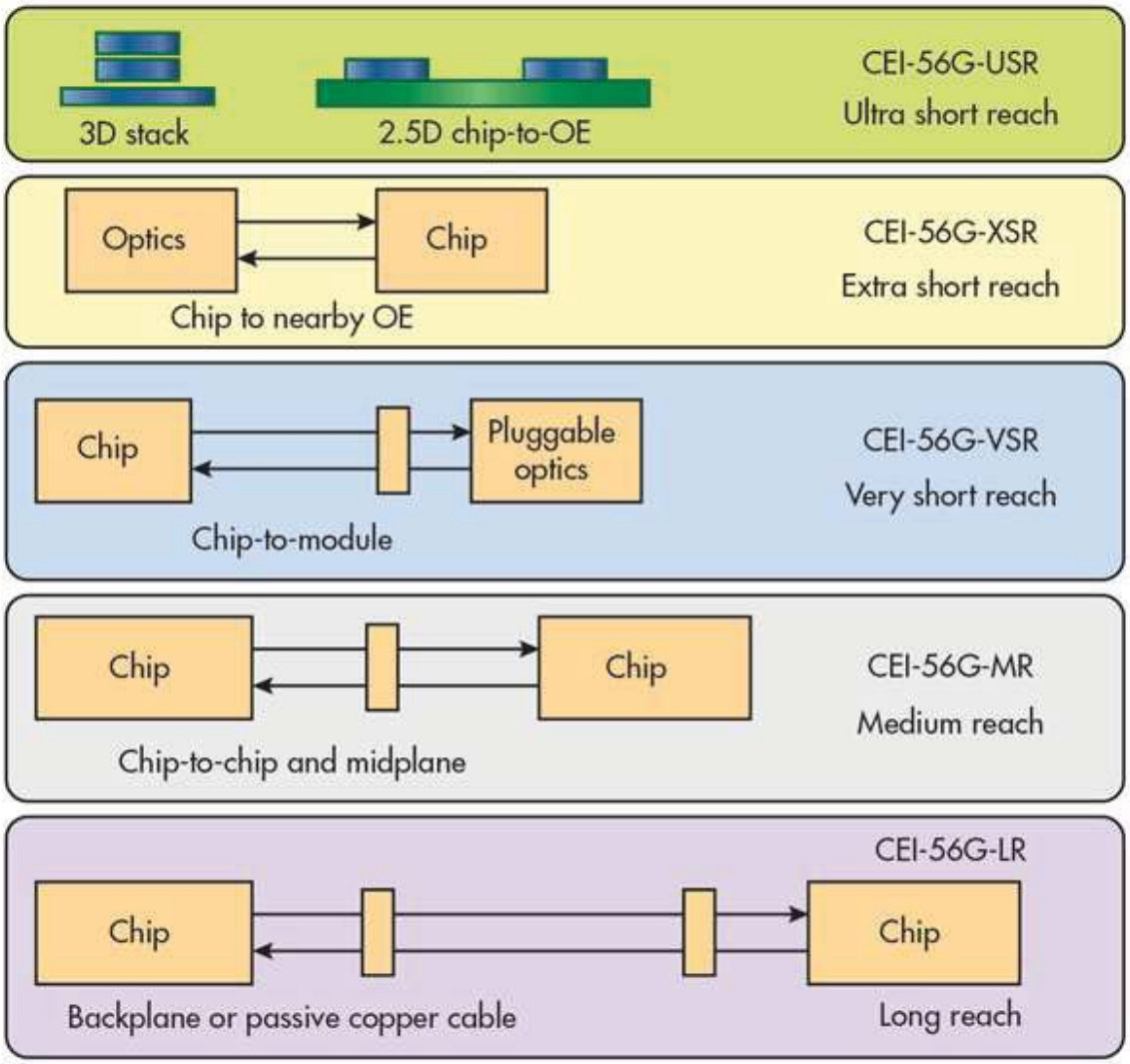
- Several standards for various media and reach

Key Changes in 50+Gb/s Standards



- Transition from NRZ to PAM4
- Relaxation of uncoded BER
  - FEC required to guarantee 1e-15 link BER
- New linearity metrics (RLM, TDECQ)
- Accommodation of ADC+DSP-based transceivers
  - E.g. Reduction in JTOL corner frequency

Common Electrical I/O (CEI) Standards



1 cm, no connectors, no packages

- 5 cm, no connectors
- 5-10 dB loss @28 GHz

- 15 cm, 1 connector
- 10 dB loss @14 GHz

- 50 cm, 1 connector
- 15-25 dB loss @14 GHz

- 100 cm, 2 connector
- 35 dB loss @14 GHz

IEEE Ethernet Standards

	Electrical Interface	Backplane	Twinax Cable	Twisted Pair (1 Pair)	Twisted Pair (4 Pair)	MMF	500m PSM4	2km SMF	10km SMF	20km SMF	40km SMF	80km SMF
10BASE-		T1S		T1S/T1L								
100BASE-				T1								
1000BASE-				T1	T							
2.5GBASE-		KX		T1	T							
5GBASE-		KR		T1	T							
10GBASE-				T1	T				BIDI Access	BIDI Access	BIDI Access	
25GBASE-	25GAUI	KR	CR/CR-S		T	SR			LR/ EPON/ BIDI Access	EPON/ BIDI Access	ER/ BIDI Access	
40GBASE-	XLAUI	KR4	CR4		T	SR4/eSR4	PSM4	FR	LR4			
50GBASE-	LAUI-2/50GAUI-2 50GAUI-1	KR	CR			SR		FR	EPON/ BIDI Access LR	EPON/ BIDI Access	BIDI Access ER	
100GBASE-	CAUI-10 CAUI-4/100GAUI-4 100GAUI-2 100GAUI-1	KR4 KR2 KR1	CR10 CR4 CR2 CR1			SR10 SR4 SR2	PSM4 DR	10X10 CWDM4/ CLR4 100G-FR	LR4/ 4WDM-10 100G-LR	4WDM-20	ER4/ 4WDM-40	ZR
200GBASE-	200GAUI-4 200GAUI-2	KR4 KR2	CR4 CR2			SR4	DR4	FR4	LR4		ER4	
400GBASE-	400GAUI-16 400GAUI- 8 400GAUI-4	KR4	CR4			SR16 SR8/SR4.2	DR4	FR8 400G-FR4	LR8 400G-LR4		ER8	ZR

50+G rates

From Ref [7]

Gray Text = IEEE Standard Red Text = In Standardization Green Text = In Study Group

Blue Text = Non-IEEE standard but complies to IEEE electrical interfaces

ethernet alliance

Standards Nomenclature

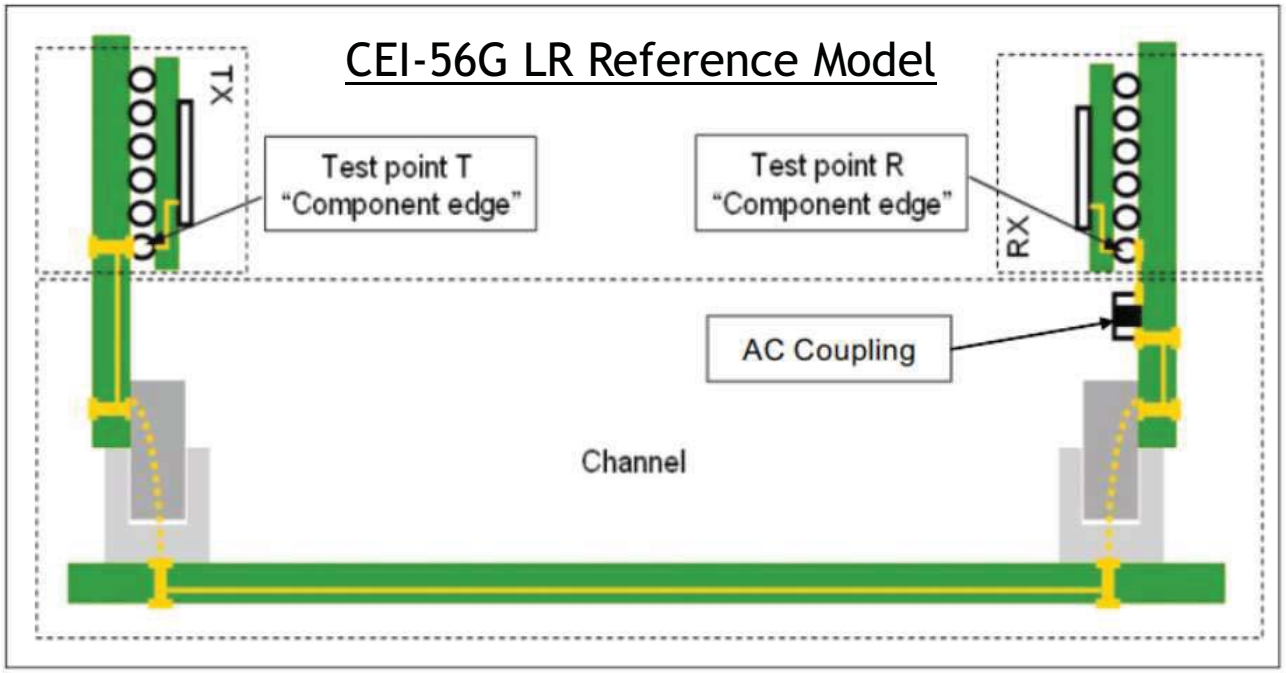
- Ethernet nomenclature: [R][mTYPE]-[L][C][n]
  - R: data rate
  - mTYPE: BASE -> Baseband
  - L: Medium/wavelength/reach
    - C-twinax Cu, D-PSM(500m), F-2km fiber, K-backplane, S-850nm, L-1310nm
  - C:PCS coding (R-64B/66B)
  - N:Number of lanes (default is 1)
- Correspondence between OIF-CEI and Ethernet standards:
  - 400GAUI-8 chip-to-module spec ↔ CEI-56G VSR
  - 200GBASE-KR4 ↔ CEI-56G LR

Outline

- Introduction
- 56G,112G Standards
  - Background
  - Electrical standards
  - Optical standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

Electrical Link Standards

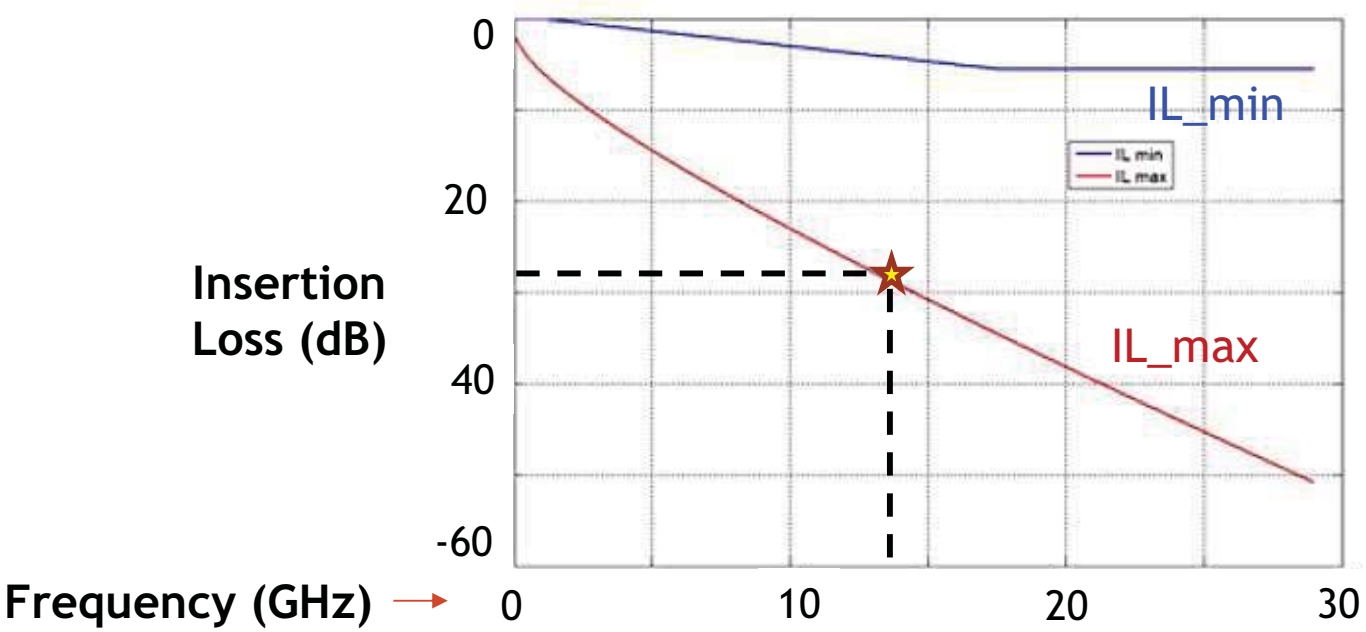
- **OIF-CEI-56G-LR** specs used as reference in following slides



Ref [5]

- Signaling spec: 28 GBaud PAM4
- BER spec: < 1e-4 (pre-FEC)

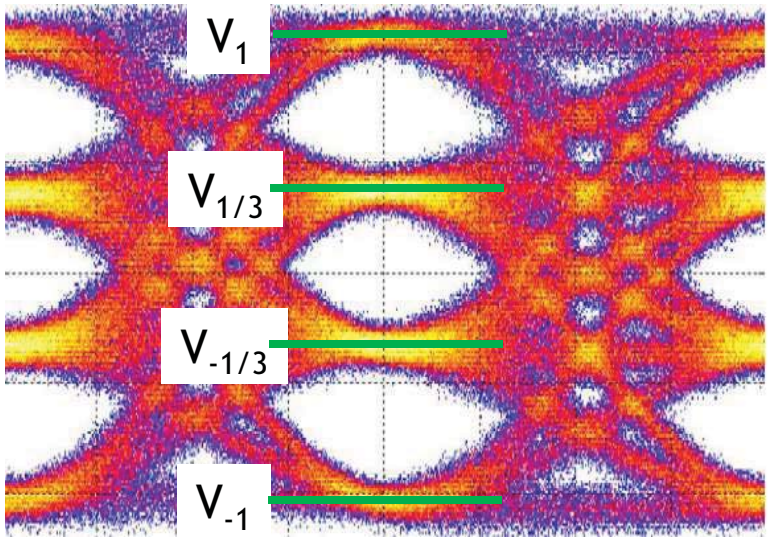
Channel Insertion Loss (IL) Spec



- Target IL < 28dB at 14 GHz (1m PCB + 2 connectors)
  - Package can add significant loss. ~35-40dB usually targeted for LR SerDes design
- Informative only. Normative spec is COM spec.
  - More on COM later.

TX Electrical Specifications: Swing, RLM

- Output swing: 0.8-1.2Vppd (no TX pre-emphasis)
- Linearity: Ratio Level Mismatch (RLM) > 0.95



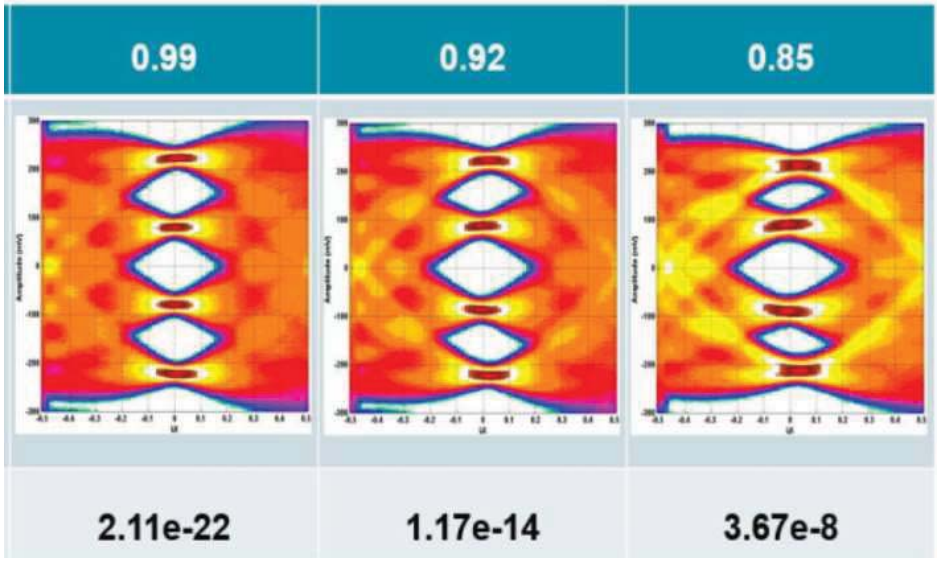
$$V_{mid} = (V_{-1} + V_{+1}) / 2$$

$$ES_1 = (V_{-1/3} - V_{mid}) / (V_{-1} - Vmid)$$

$$ES_2 = (V_{+1/3} - V_{mid}) / (V_{+1} - Vmid)$$

$$R_{LM} = \min((3 \cdot ES_1), (3 \cdot ES_2), (2 - 3 \cdot ES_1), (2 - 3 \cdot ES_2))$$

Effect of RLM on BER



Ref [8]

TX Electrical Specifications: SNDR

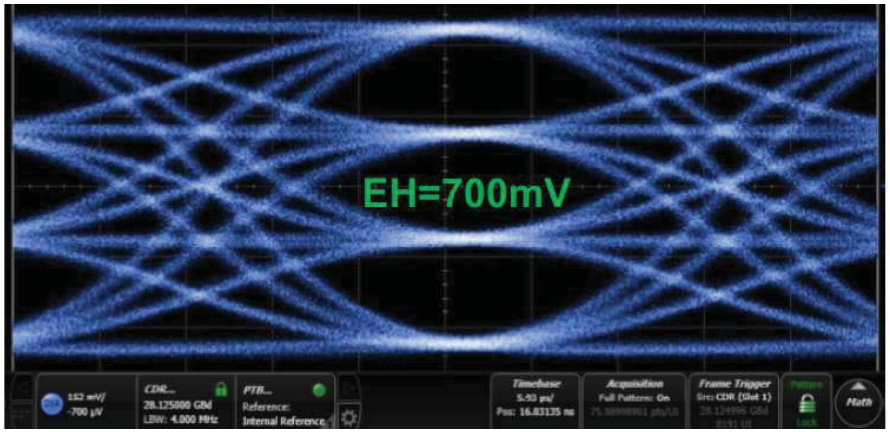
- SNDR spec is intended to constrain distortion and uncorrelated noise at TX output (discounts ISI)
- Computed from TX output and linear fit pulse response p(k)

Max(p(k))

$$SNDR = 10\log_{10}\left(\frac{p_{max}^2}{\sigma_e^2 + \sigma_n^2}\right)(dB)$$

Distortion Random noise

Measured 56G TX Eye [9]

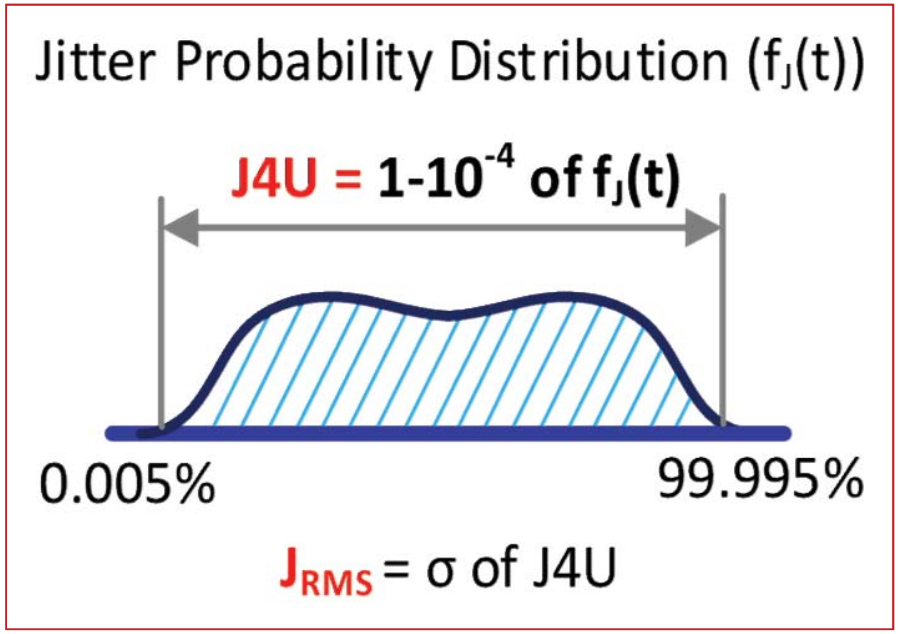


RLM = 0.99, SNDR = 37 dB

- Require SNDR > 31 dB

TX Electrical Specifications: Jitter

- 3 jitter components computed from 12 possible PAM4 transitions
  - J_{RMS}, J4U, EOJ
  - EOJ = duty-cycle distortion

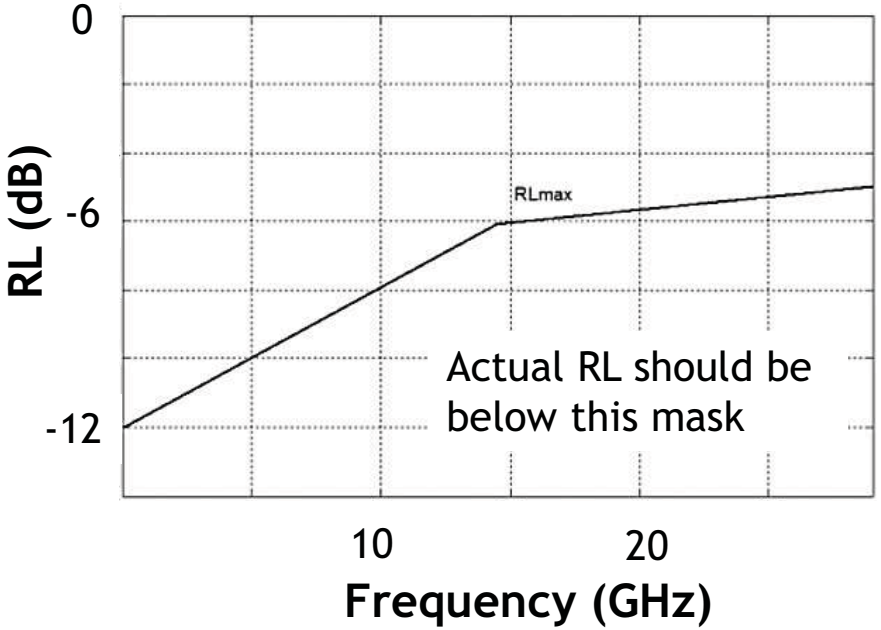


- Jitter sources:
  - Random: Thermal/flicker noise from PLL, clock distribution
  - Systematic: DCD/quadrature error, bandwidth limits, supply noise

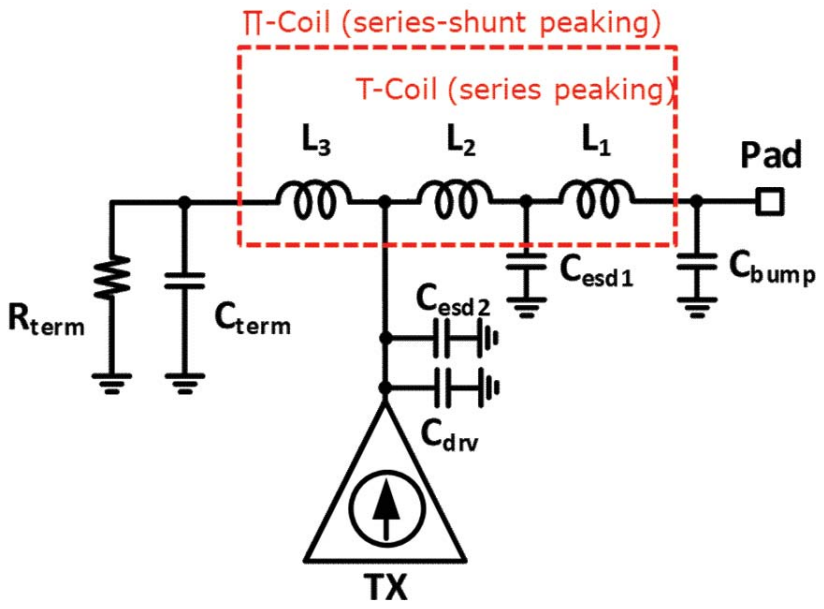
Jitter	Max Value
J_{RMS}	0.023 UI
J4u	0.118 UI
EOJ	0.019 UI

TX Electrical Specifications: Return Loss

- Intended to limit impedance discontinuity at TX



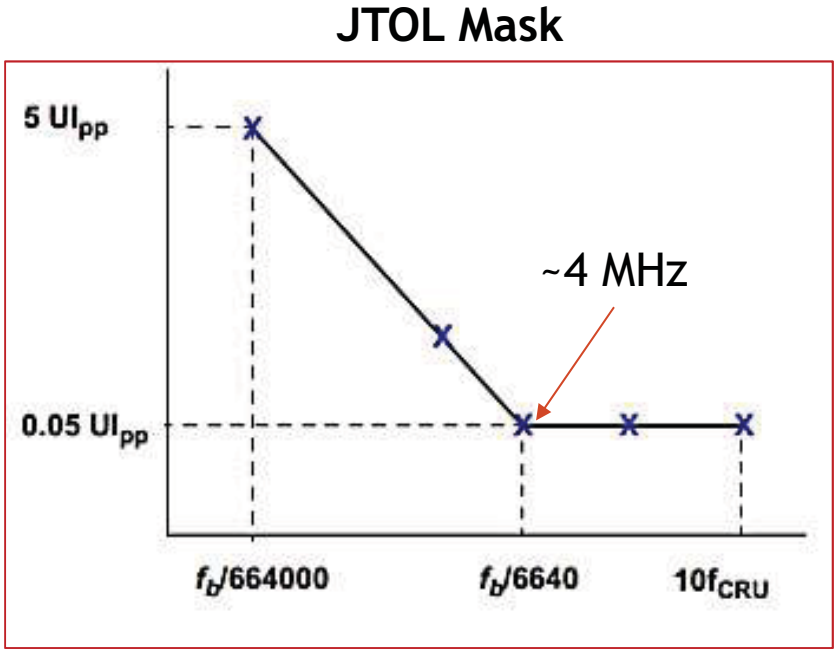
Example pad network design to improve RL [26]



- Meeting RL spec requires careful design of pad network

RX Electrical Specifications

- RX should meet target BER with worst-case compliant TX + channel
- Two stress tolerance tests:
  - Interference tolerance
  - Jitter tolerance (JTOL)



- JTOL corner frequency relaxed to accommodate ADC+DSP-based RX
  - from 10MHz in CEI-28G to 4MHz in CEI-56G/112G

Outline

- Introduction
- 56G,112G Standards
  - Background
  - Electrical standards
  - Optical standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

56G/112G Optical Standards

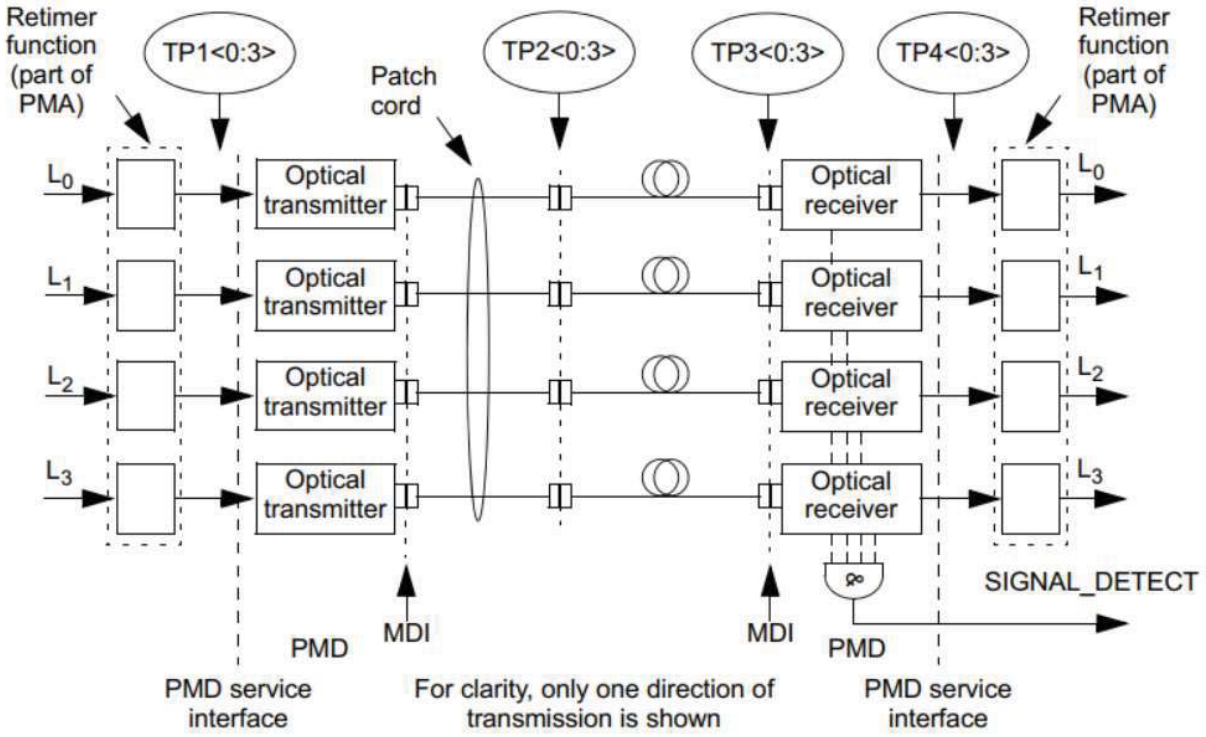
	100m SR	500 m DR	2 km FR	10 km LR
50 Gb/s	1 x 50 Gb/s	1 x 50 Gb/s	1 x 50 Gb/s	1 x 50 Gb/s
100 Gb/s	2 x 50 Gb/s Parallel 100GBASE-SR2	1 x 100 G Single lane 100GBASE-DR	N/A	N/A
200 Gb/s	4 x 50 Gb/s Parallel 200GBASE-SR4	4 x 50 Gb/s Parallel 200GBASE-DR4	4λ x 50 Gb/s CWDM 200GBASE-FR4	4λ x 50 Gb/s LWDM 200GBASE-LR4
400 Gb/s	16 x 25G Parallel 400GBASE-SR16	4 x 100 Gb/s Parallel 400GBASE-DR4	8λ x 50 Gb/s LWDM 400GBASE-FR8	8λ x 50 Gb/s LWDM 400GBASE-LR8

802.3cd (May 2016-18)
Sept'16: D1.0 spec
Nov'16: Last new proposal
May'17: D2.0
Nov'17: Last change
Jan'18: D3.0
Sept'18: Standard

802.3bs (May 2014-17)
Sept'15: D1.0 spec
July'16: D2.0
Nov'16: Last Change
March'17: D3.0
Dec'17: Standard

- Following material uses **400GBASE-DR4** as reference

400GBASE-DR4



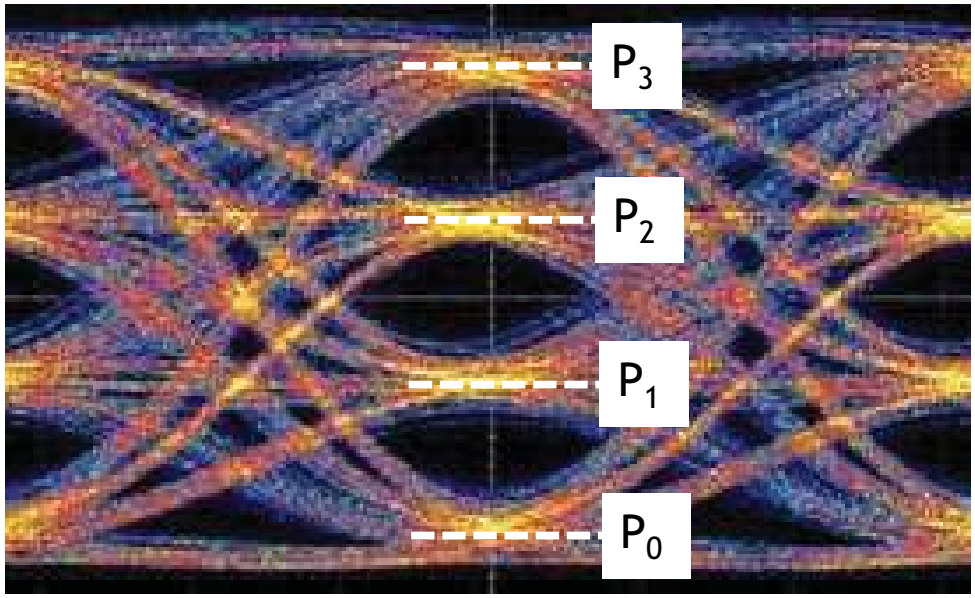
- 4 lanes of 53 GBaud PAM-4 over 500m SMF
  - Target BER < 2.4e-4 (pre-FEC)
- Spec defines TX(RX) signal characteristics at TP2(TP3)

400GBASE-DR4 TX Specs

Spec	Value
Signaling rate	53.125 GBd
Modulation	PAM4
Outer OMA	-0.8 dBm to 4.2 dBm
TDECQ (max)	3.4 dB
OMA minus TDECQ (min)	-2.2 dBm
Extinction Ratio (ER) (min)	3.5 dB

- Key specs (measured at TP2):
  - (Outer) OMA, ER
  - TDECQ (NEW)
- No mask test like previous NRZ specs

PAM4 OMA, ER Definition



- Outer OMA = (P3 – P0)
- Outer ER = (P3 / P0)

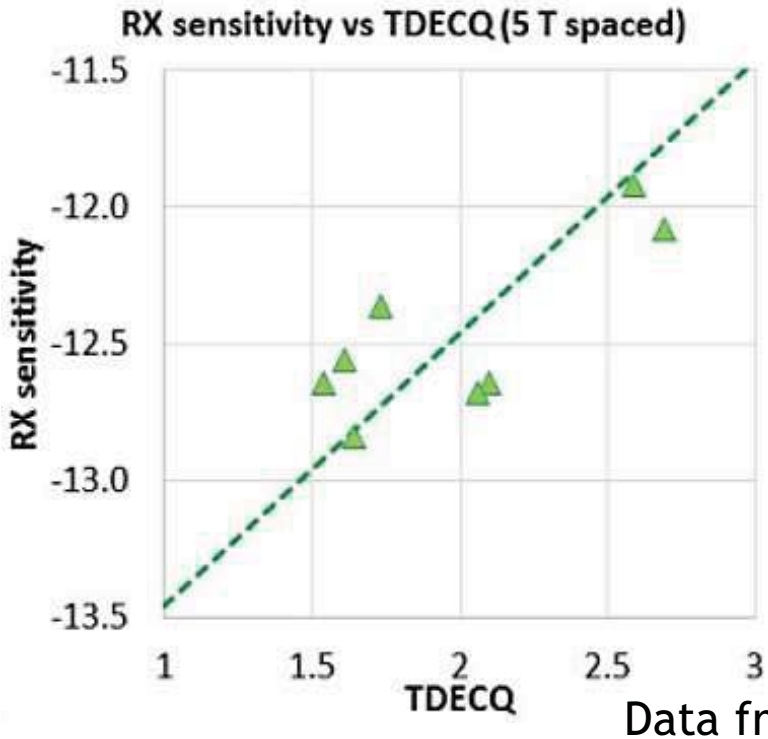
} 10*log10() to convert to dB

- Average power can be computed from OMA, ER
  - P_{avg} = (OMA * (ER + 1)/(2 * (ER - 1))

TDECQ Metric

- TDECQ = TX and Dispersion Eye Closure Quaternary
  - Replaces TDP and mask tests in previous NRZ specs
- Intended to estimate power penalty due to TX and path impairments

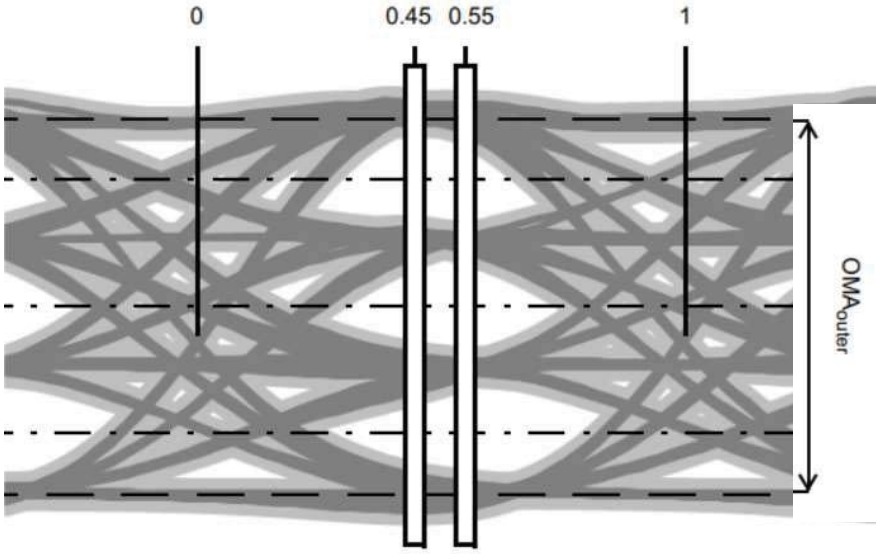
Good correlation observed between TDECQ and RX sensitivity



Data from [10]

TDECQ Definition

- A measure of noise margin of actual TX versus ideal TX for a worst-case channel + reference RX



$$\text{TDECQ} = 10\ log_{10}\left[\frac{OMA_{outer}}{6} \times \frac{1}{Qt \cdot R}\right]$$

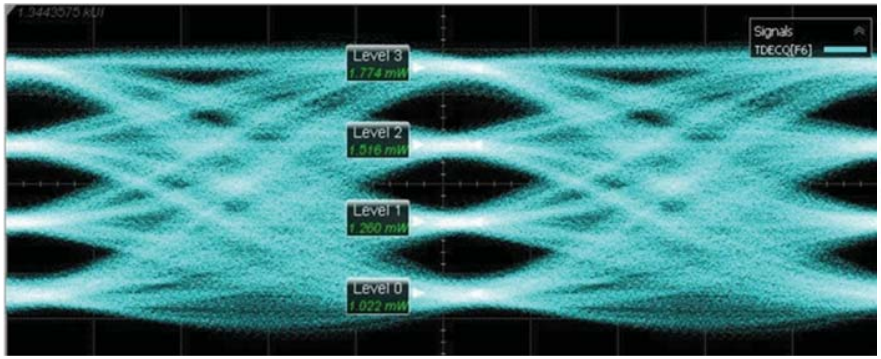
with Q_t as Q-function consistent with target BER (Q_t = 3.414 for 2.4e-4 BER)
with R as RMS noise term of the receiver

- Reference receiver parameters:
  - 4th order BT-filter with (baud rate / 2) bandwidth
  - 5-tap symbol-spaced FFE

Example TDECQ Measurements

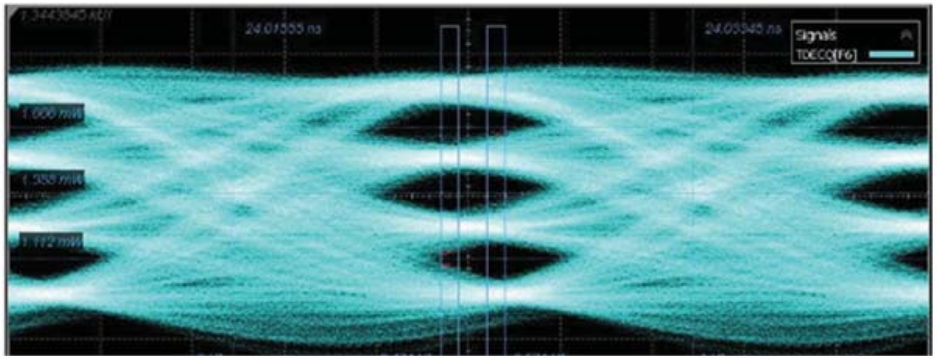
112G Optical TX Output for 2 (driver + modulator) settings [11]

With nonlinearity compensation



0.5dB TDECQ

Without nonlinearity compensation



2.0dB TDECQ

- TDECQ captures the effect of modulator nonlinearity and bandwidth on link budget

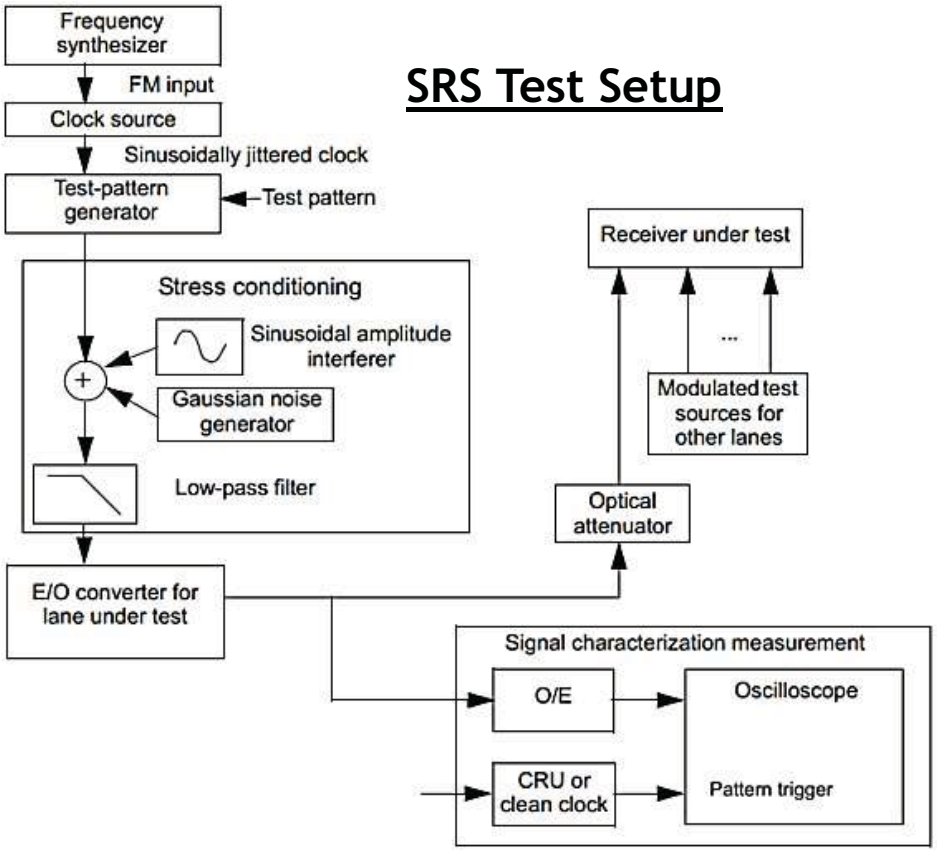
400GBASE-DR4 RX Specs

Spec	Value
Signaling rate	53.125 GBd
Modulation	PAM4
Avg RX power	-5.9 dBm to 4 dBm
RX Sensitivity (max)	-4.4 dBm
Stressed RX Sensitivity (SRS) (max)	-1.9 dBm

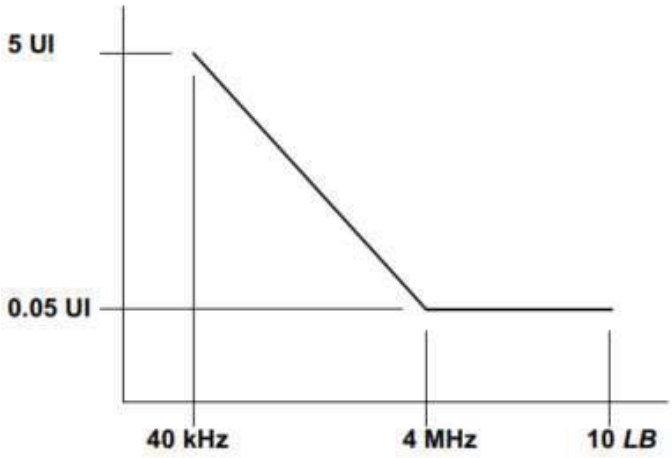
- Key specs:
  - RX sensitivity (informative),
  - SRS (normative)

Stressed RX Sensitivity (SRS) Test

SRS Test Setup

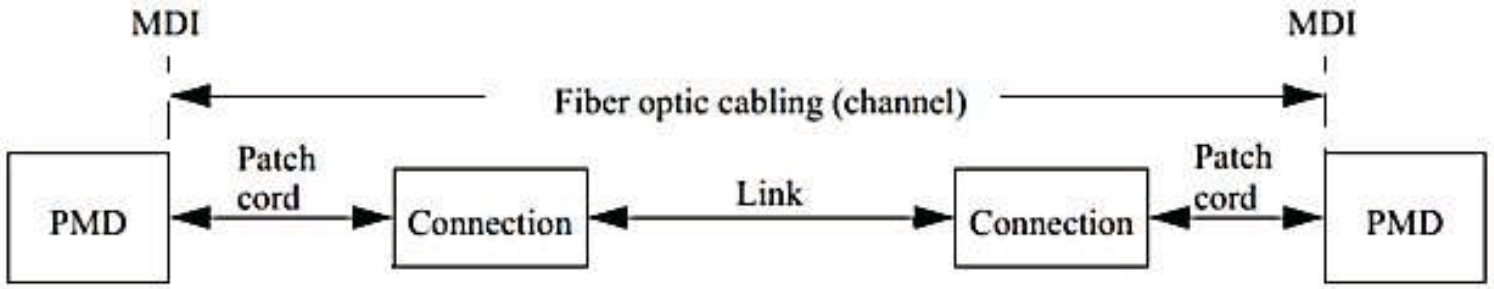


Sinusoidal JTOL Mask



- Configure stressors to mimic worst-case TX
- JTOL test (and specs) similar to that for electrical RX

Optical Channel Specs

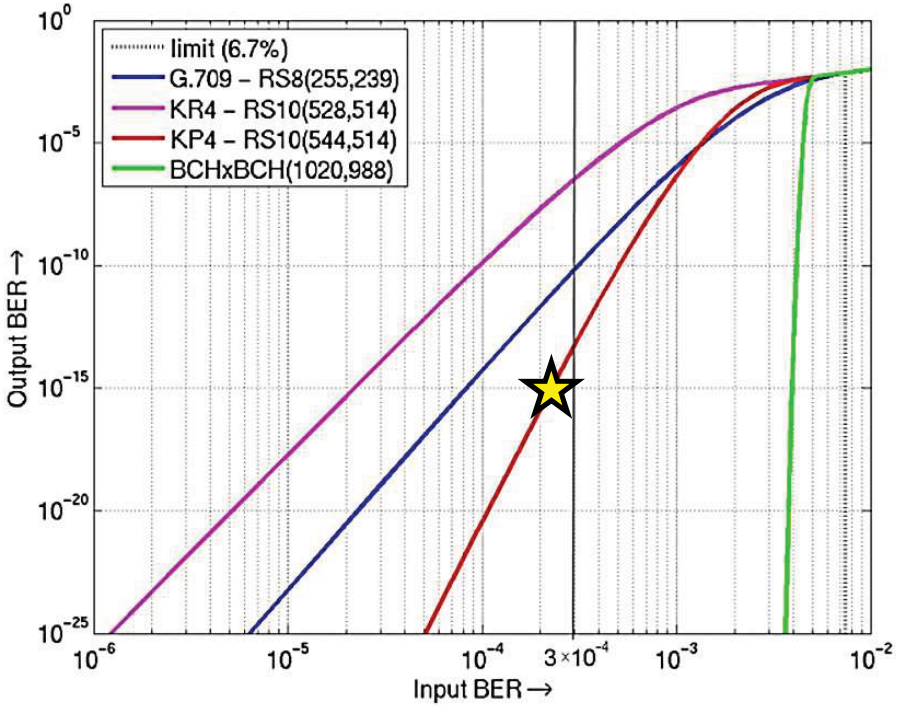


400G-DR4 Channel Characteristics

Description	400GBASE-DR4	Unit
Operating distance (max)	500	m
Channel insertion loss[a,b] (max)	3	dB
Channel insertion loss (min)	0	dB
Positive dispersion[b] (max)	0.8	ps/nm
Negative dispersion[b] (min)	–0.93	ps/nm
DGD_max[c]	2.24	ps
Optical return loss (min)	37	dB

- Fiber loss < 0.5dB/km
- 2.75dB allocated for connector and splice loss
- Fiber dispersion causes negligible penalty for <2km reach

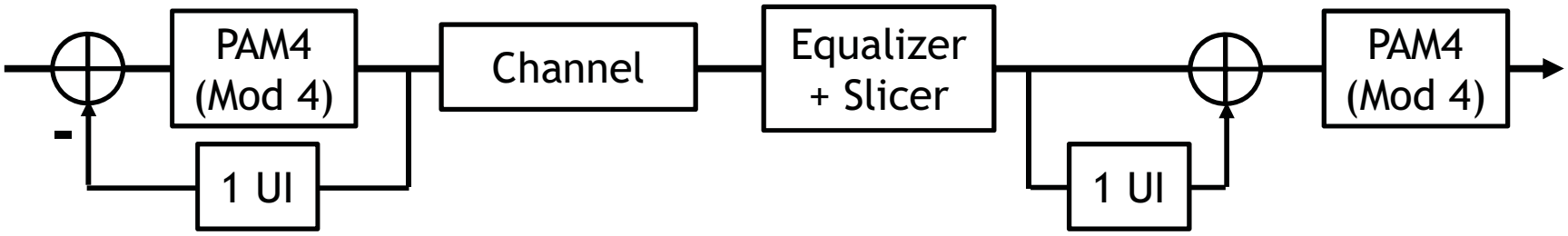
FEC in 56G/112G PAM4 Links



Code Name	Symbol size, m (bits)	Correctable symbols per block, t	Rate Overhead (%)	Random Error Coding Gain (dB)	Input BER for 10^{-15} corrected BER
G.709 RS8(255,239) [8]	8	8	6.7	5.4	$8\cdot10^{-5}$
IEEE 100GBASE-KR4 RS(528,514) [25]	10	7	2.7	4.9	$2\cdot10^{-5}$
IEEE 100GBASE-KP4 RS(544,514) [25]	10	15	5.8	6.1	$2\cdot10^{-4}$
ITU G.975.1 (I.9) [28] BCHxBCH(1020,988)	1	N/A due to concate-nated code	3.2	9	$4\cdot10^{-3}$

- FEC designed to improve BER from 10^{-4} to 10^{-15}
  - KP4 FEC RS(544,514) is commonly used
- Burst error length should not exceed code symbol size
  - 10bits for KP4 FEC

Pre-coding to Limit DFE Error Propagation



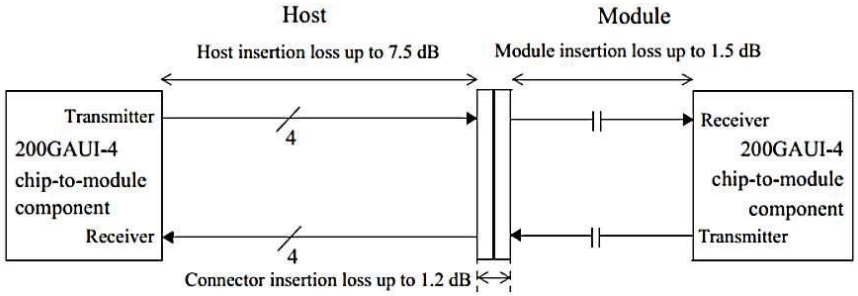
- 1/(1+D) pre-coding converts DFE burst error into two symbol errors [12]
  - Also doubles error rate for random errors
- Can reduce loss in coding gain for large 1st DFE tap weight
  - Effectiveness analyzed in detail in [13]

Outline

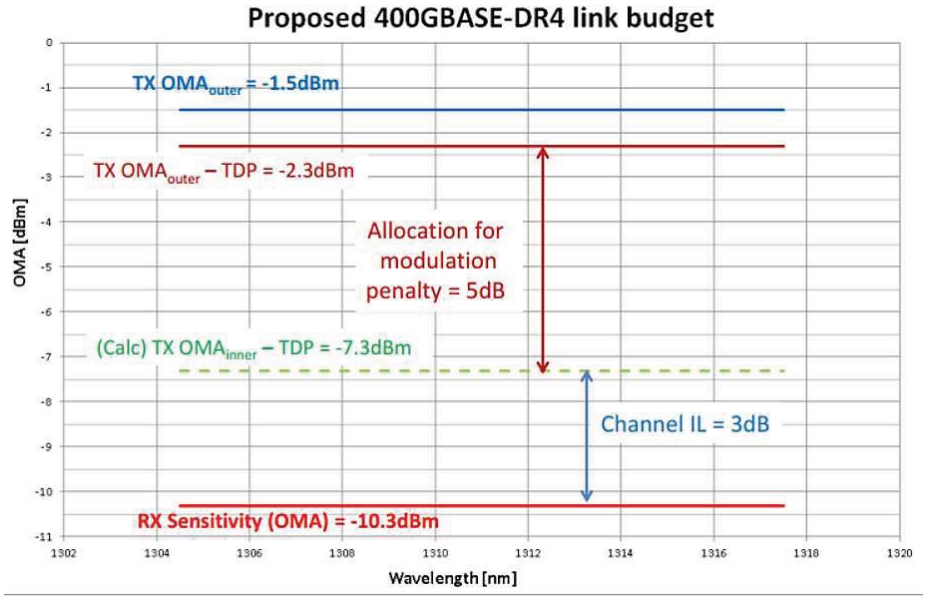
- Introduction
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

Link Budgeting: Objective

200GAUI-4 C2M IL budget



Proposed 400GBASE-DR4 link budget

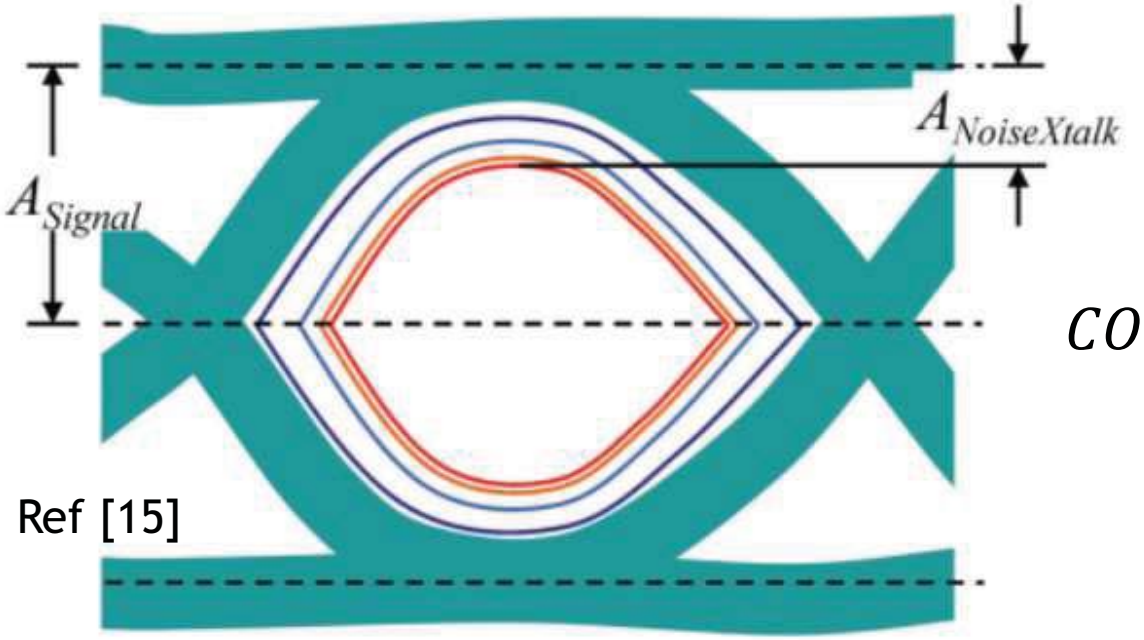


- Specify tolerable channel and device impairments to meet standards targets
- Channel Operating Margin (COM) is a standardized way to do this for electrical links

Channel Operating Margin (COM)

- A unified budget that ties TX, RX and channel specifications together [14]
- Replaces frequency-domain mask-based channel specifications
  - For IL, ILD, crosstalk etc.
- Used extensively in IEEE and OIF standards process:
  - To evaluate proposals
  - For balancing PHY and interconnect interests

COM Definition



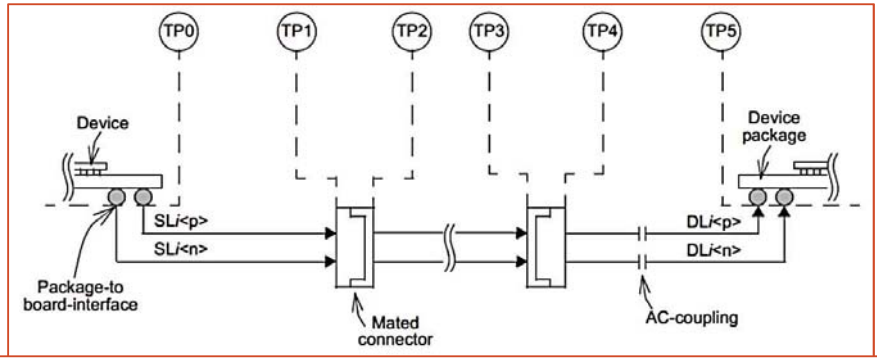
Ref [15]

$$COM = 20 \cdot log \frac{A_{Signal}}{A_{NoiseXtalk}}$$

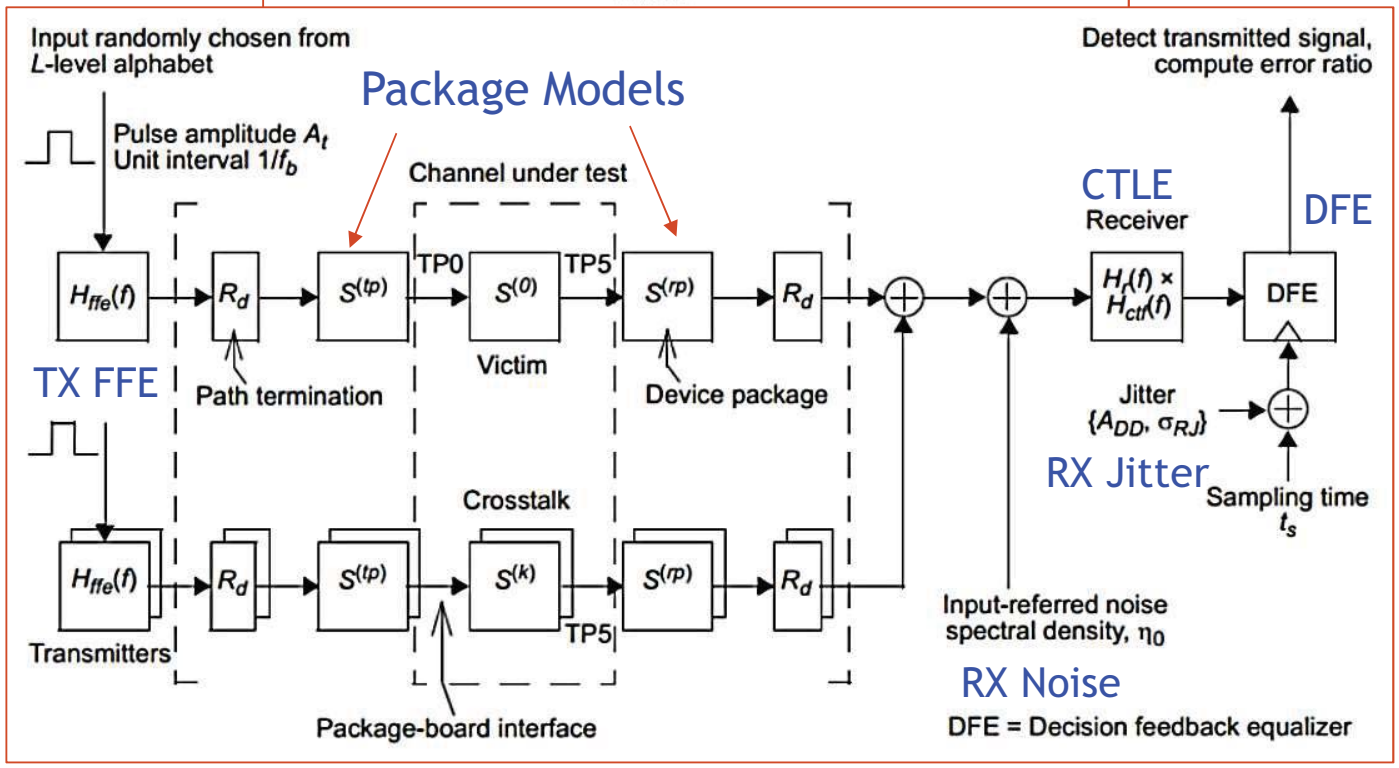
- COM is an SNR metric
  - Computed using a reference model
- Most standards require COM > 3 dB at target BER

COM Reference Model

Physical Topology



Analytical Model



COM Computation - Step 1 (SBR)

- Compute single-bit response (SBR)
  - Cascade channel component S-parameters and EQ filters

- **Cascaded channel**

$$H^{(k)}(f) = H_{ffe}(f)H_{21}^{(k)}(f)H_r(f)H_{ctf}(f)$$

- **TX FFE**

$$H_{ffe}(f) = \sum_{i=-1}^{1} c(i)\exp(-j2\pi(i+1)(f/f_b))$$

- **Passive link**

$$H_{21}(f) = \frac{s_{21}(f)(1-\Gamma_1)(1+\Gamma_2)}{1-s_{11}(f)\Gamma_1(f)-s_{22}(f)\Gamma_2(f)+\Gamma_1(f)\Gamma_2(f)\Delta S(f)}$$

- **RX CTLE equalizer**

$$H_{CTLE}(f) = f_b \frac{j\cdot f + 0.25\cdot f_b 10^{\frac{G_{DC}}{20}}}{(j\cdot f + 0.25\cdot f_b)\cdot(j\cdot f + f_b)}$$

RX filter usually a 4th order BT-filter with defined BW

COM Computation - Step 2 (EQ Search)

- Determine ‘optimal’ equalization settings
- Simplified FOM used to make FFE+CTLE search computationally efficient

$$FOM = 10\log_{10}\left(\frac{A_s^2}{\sigma_{TX}^2 + \sigma_{ISI}^2 + \sigma_J^2 + \sigma_{XT}^2 + \sigma_N^2}\right)$$

From TX SNDR

Channel ISI

Jitter

Crosstalk

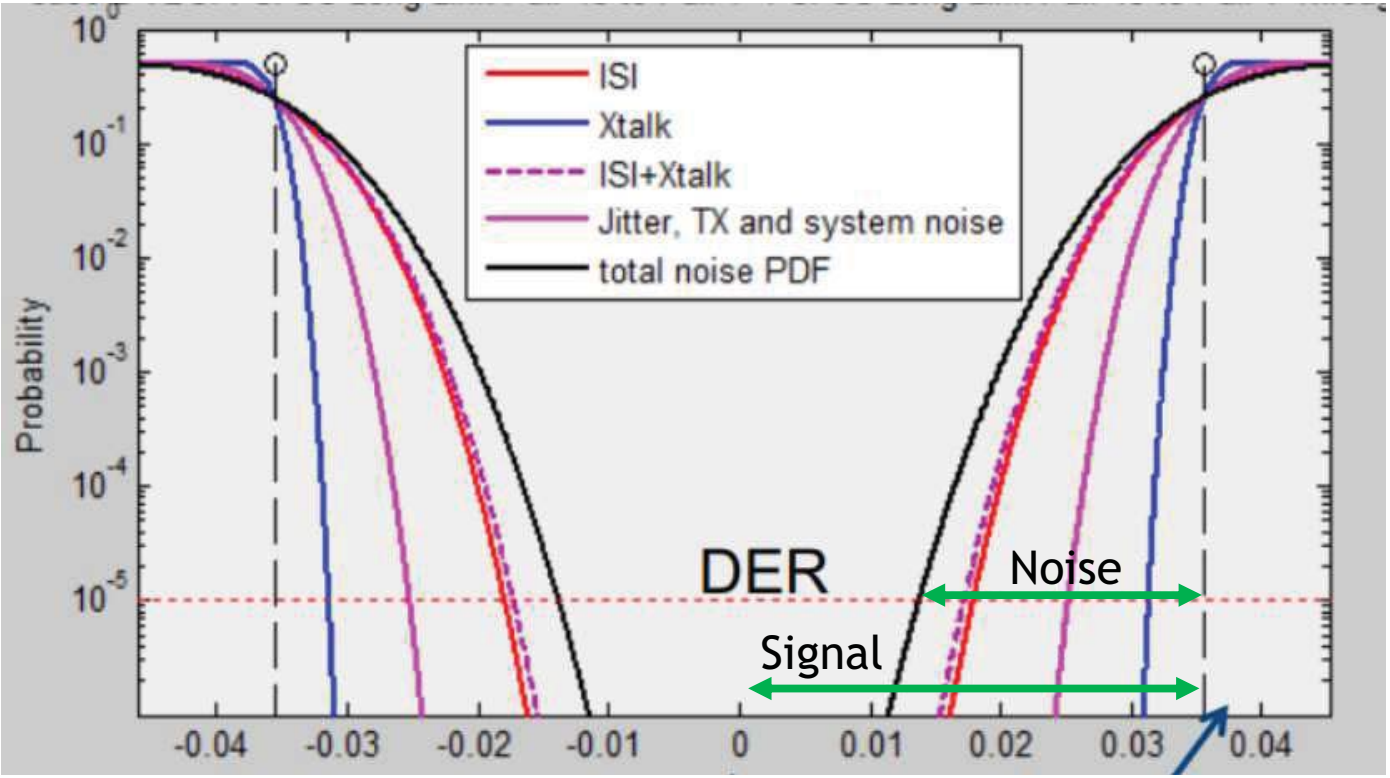
Noise at CTLE output

- For details, see [14]

COM Computation – Step 3

- Use optimal equalizer parameters to compute interference and noise probability density functions (PDFs)
  - Interference: ISI, Xtalk
  - Noise: RX noise, jitter-induced voltage noise
- Convolve these PDFs and integrate to compute noise amplitude for COM calculation

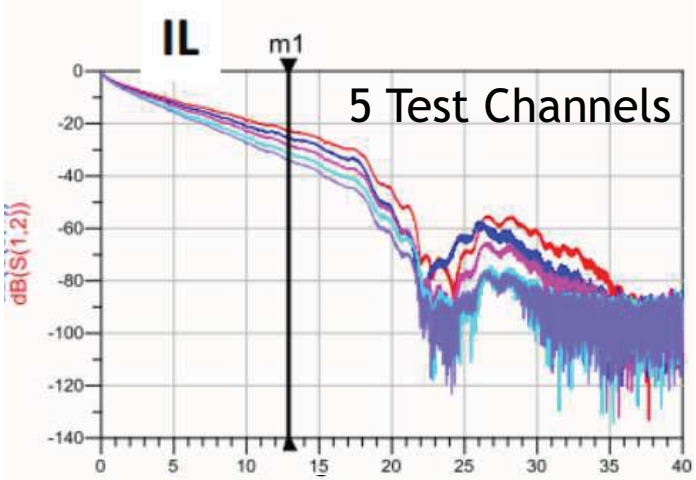
Example Result

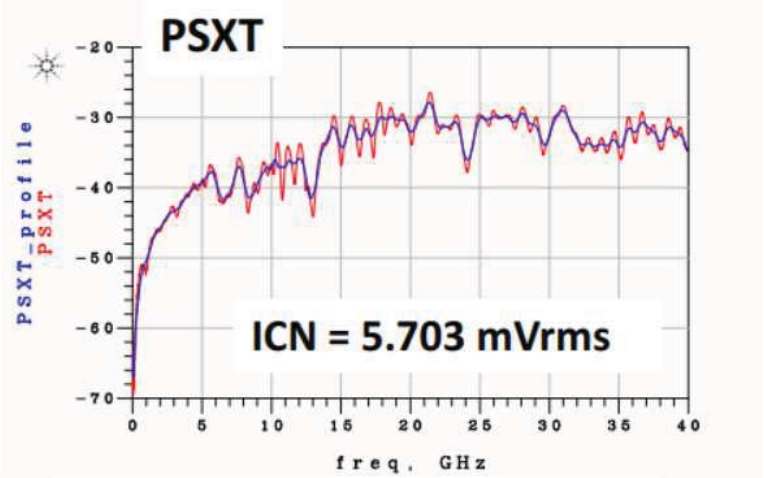


Ref [14]

- Signal = 35 mV
- Noise at DER = 35 – 13 = 22 mV
- COM ~ 4dB > 3 dB target

Example Usage for Channel Budgeting





COM (dB)	Link #1	Link #2	Link #3	Link #4	Link #5
Link Loss	22.8	25.6	28.2	31.5	33.9
w/o Crosstalk	5.00	4.32	3.99	3.18	2.55
w/ Crosstalk	-0.86	-2.27	-3.20	-4.94	-5.95
Degradation	5.86	6.59	7.19	8.12	8.50

- Compare 5 channels with different IL profiles
- COM analysis shows:
  - importance of crosstalk
  - Limits of chosen SerDes EQ capability

Example Usage for Equalizer Budgeting

COM Results with 112G Test Channels

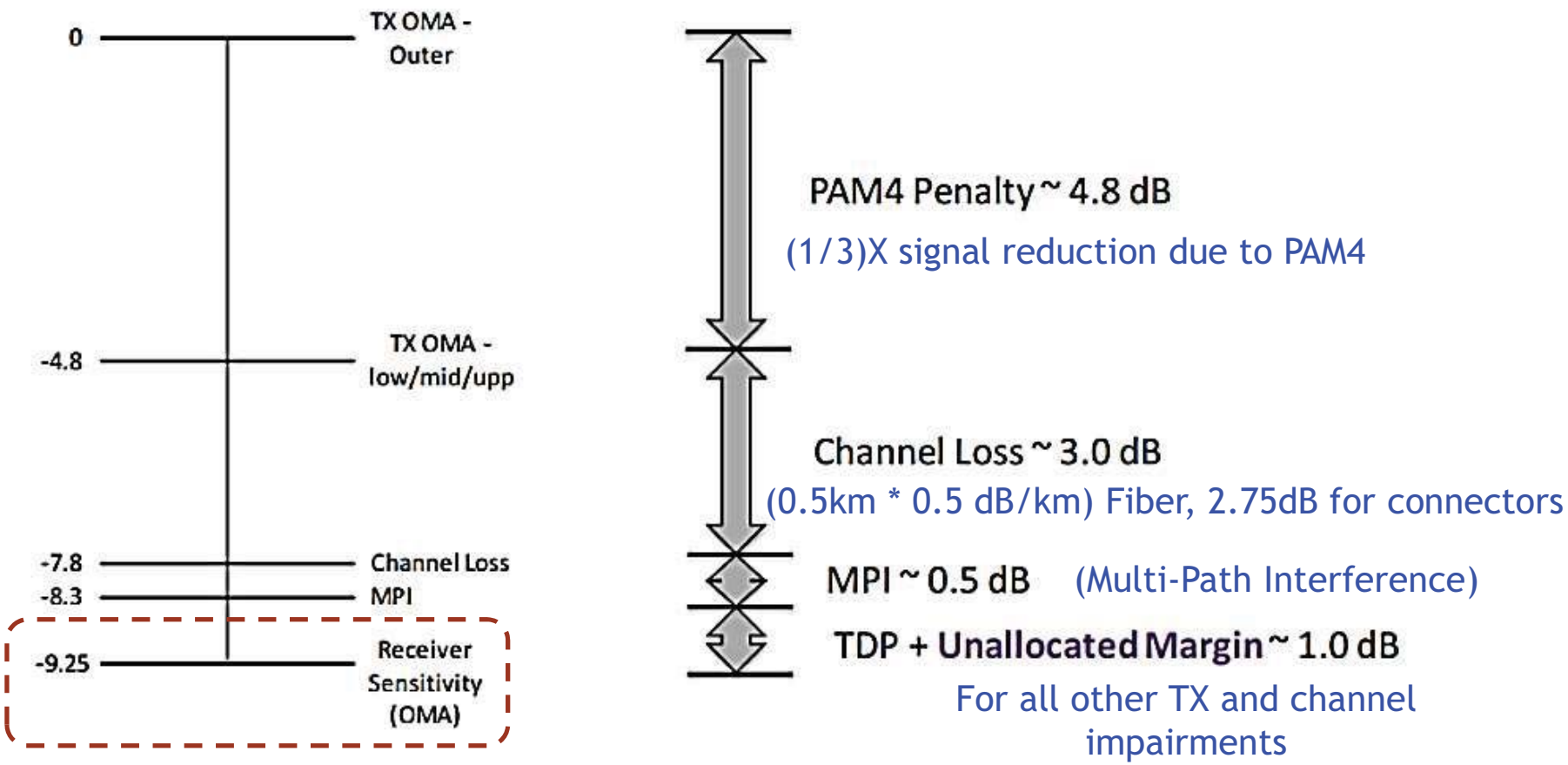
EQ:1-tap DFE + FFE (3 pre-/N-post) Ref [16]

Vendor	Channel w/o XTLK	FFE Tap Length (Post-tap Length)							
		16 (12)	17 (13)	18 (14)	19 (15)	20 (16)	24 (20)	28 (24)	32 (28)
Intel	BP_2conn_85ohm_30dB_Nom_t	-1.93	-1.88	-0.64	0.44	0.62	1.72	1.94	1.99
Intel	BP_2conn_85ohm_30dB_HzLzHz_t	-2.29	-2.22	-1.15	0.1	0.45	2.11	2.38	2.43
Intel	BP_2conn_85ohm_30dB_LzHzLz_t	-1.53	-1.46	0.16	0.98	0.99	2.33	2.63	2.69
Samtec	BP__Z100sm_IL15to16_BC-BOR_N_N_N_t	-0.58	-0.53	0.6	1.82	1.88	2.45	3.01	3.07
Sametc	BP__Z100sm_IL25_27_BC-BOR_N_N_N_t	-0.56	-0.47	0.22	1.51	1.57	2.05	2.55	2.79
Sametc	BP__Z100sm_IL30to32_BC-BOR_N_N_N_t	-1.09	-0.93	-0.64	0.55	0.56	0.96	1.39	1.56
Samtec	CAd2d__2p0m_awg28_m_BC-BOR_N_N_N_t	-0.06	-0.06	0.84	2.17	2.17	2.19	2.2	2.21
Samtec	CAd2d__2p5m_awg28_m_BC-BOR_N_N_N_t	-0.32	-0.21	0.41	1.65	1.65	1.69	1.69	1.69
TE	G1112_Ortho_t	1.65	1.57	3.05	5.42	5.42	5.43	5.61	5.75
TE	B56_CblBP_t	0.29	0.21	1.78	4.29	4.29	4.36	4.45	4.47

- COM analysis reveals DSP complexity required to accommodate 1-tap DFE for target channels

Illustrative Optical Link Budget

Proposal for 400GBASE-DR4

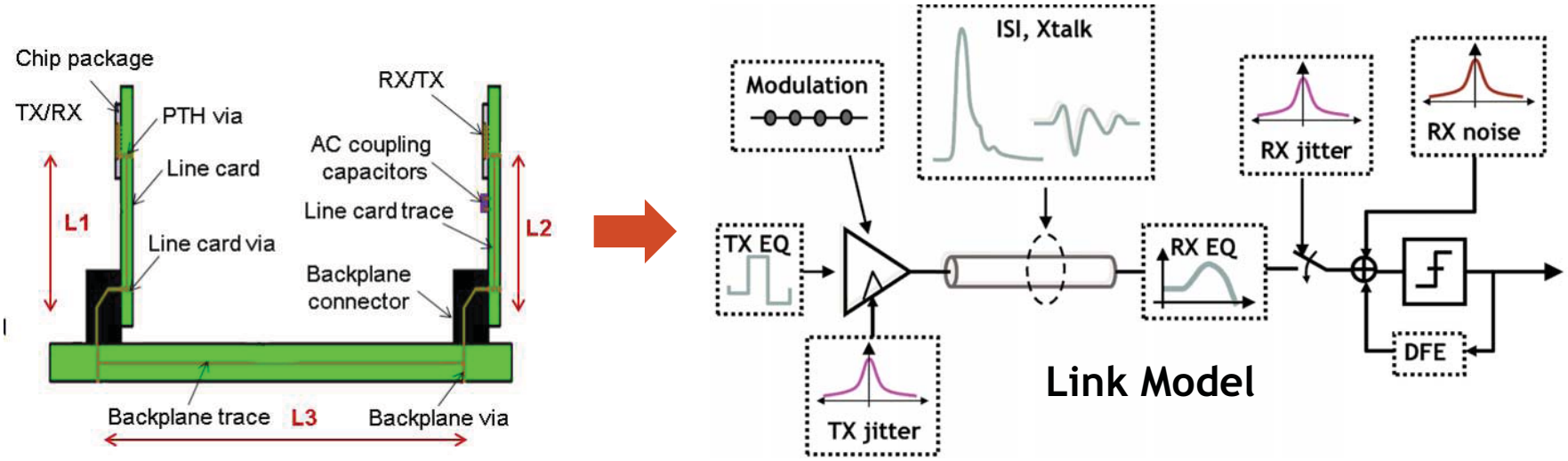


Minimum required RX sensitivity for this budget

Outline

- Introduction
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
  - Modeling Methods
  - Electrical Links
  - Clocking
  - Optical Links
- Conclusion

Link Modeling: Objectives



- Link architecture definition:
  - E.g. TX/RX equalization, clock recovery method etc.
- Component specification and sensitivity analysis:
  - E.g. PLL jitter, ADC/DAC resolution, RX noise etc.
- Link validation

Link Modeling: Methods

- Statistical analysis:
  - Probabilistic analysis assuming LTI system [17-22]
  - E.g. StatEye, JnEye (Intel), LinkLab (Rambus) …
- Time-domain simulation:
  - Based on behavioral models of link components
  - Uses Matlab, Simulink, Verilog/VerilogA …
- Event-driven simulation:
  - Uses functional forms of analog waveforms to break accuracy-time step tradeoff [23]
- IBIS-AMI:
  - Industry standard to describe link components [24]
  - Used primarily for link validation

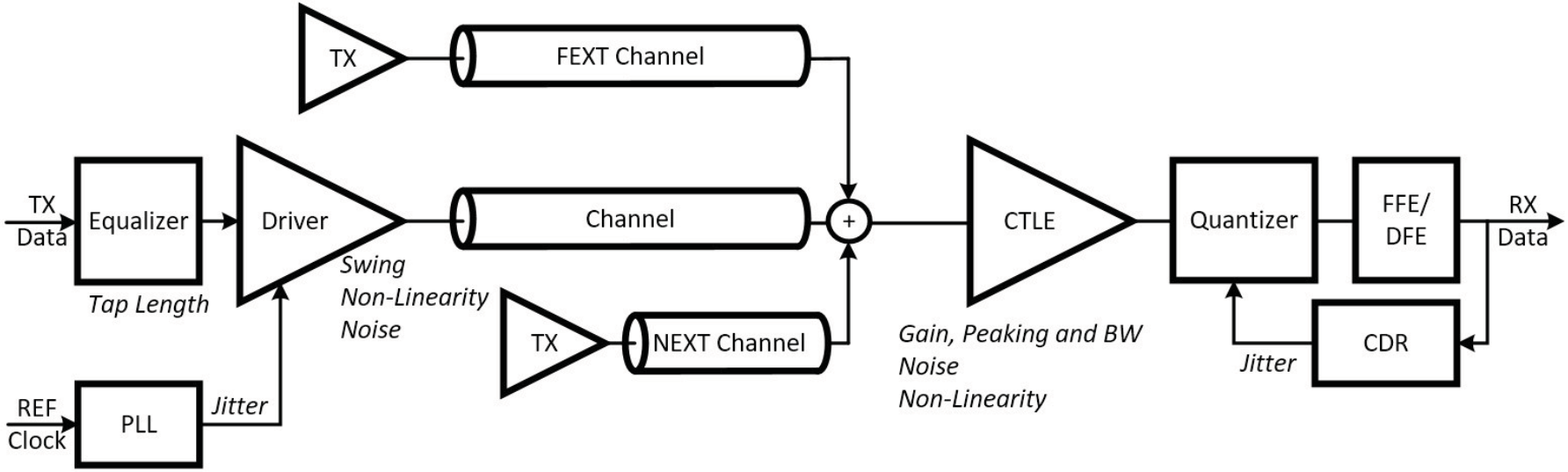
Time-Domain Link Modeling & Simulation

- Most commonly used for 56G/112G electrical & optical link design
- Target pre-FEC BER of 1e-4 to 1e-6 makes this method practical
- Can accommodate non-LTI behavior, quantization effects and calibration loops

Outline

- Introduction
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
  - Modeling Methods
  - Electrical Links
  - Clocking
  - Optical Links
- Conclusion

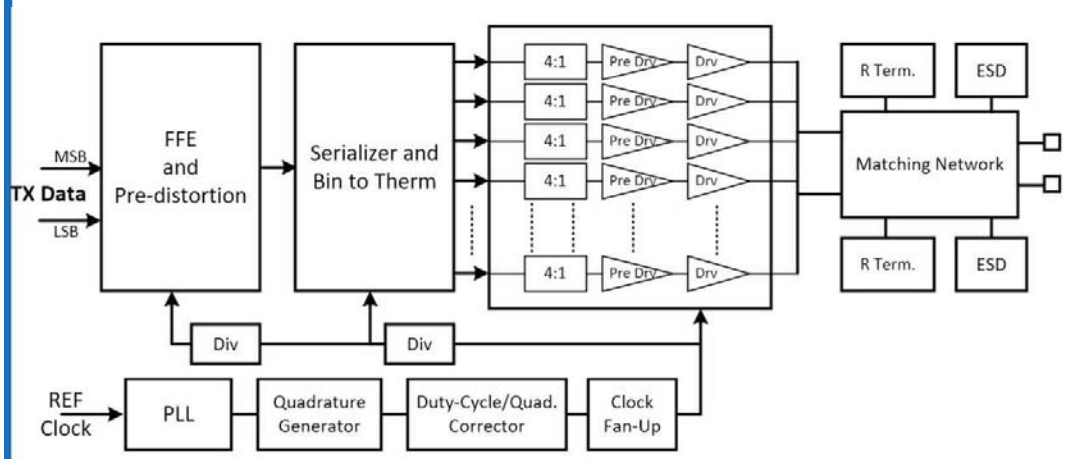
Electrical Link Model



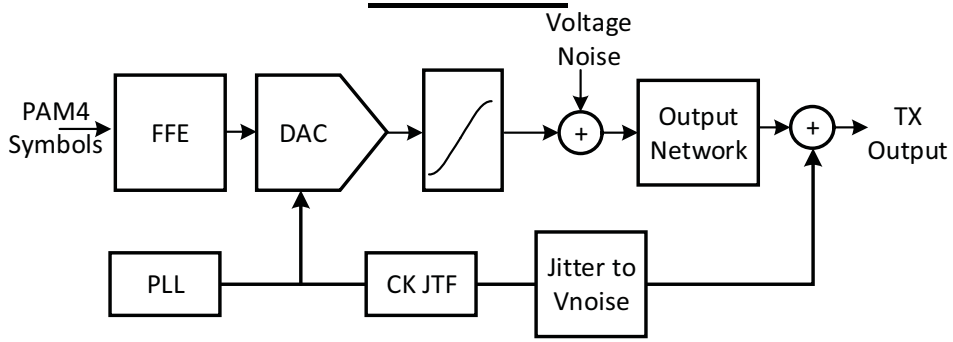
- Mix of frequency and time domain models
  - Frequency domain: Channel models, matching networks, analog filters
  - Time domain: Clocks, ADC, DSP …
- Crosstalk channels are typically asynchronous with victim

Electrical TX Model

Example 112G TX block diagram



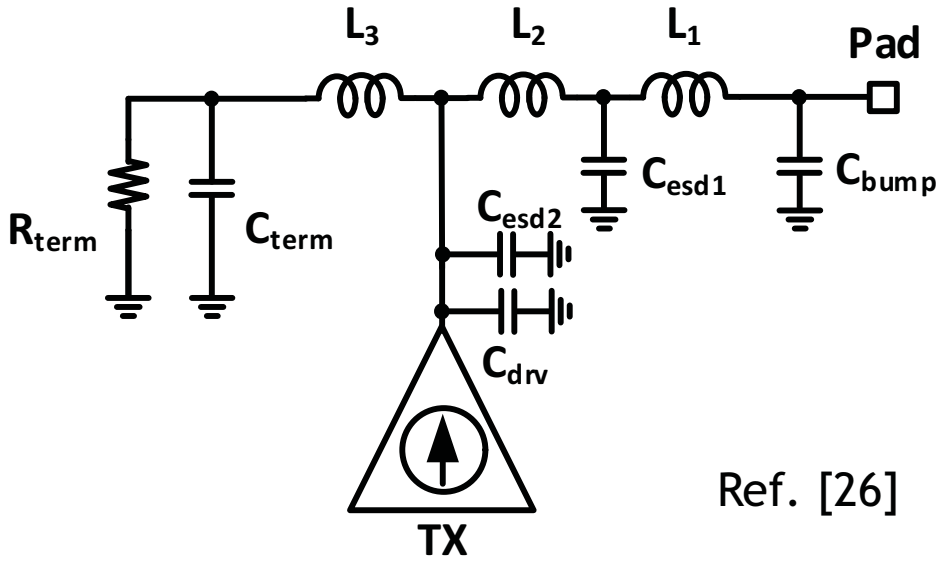
TX Model



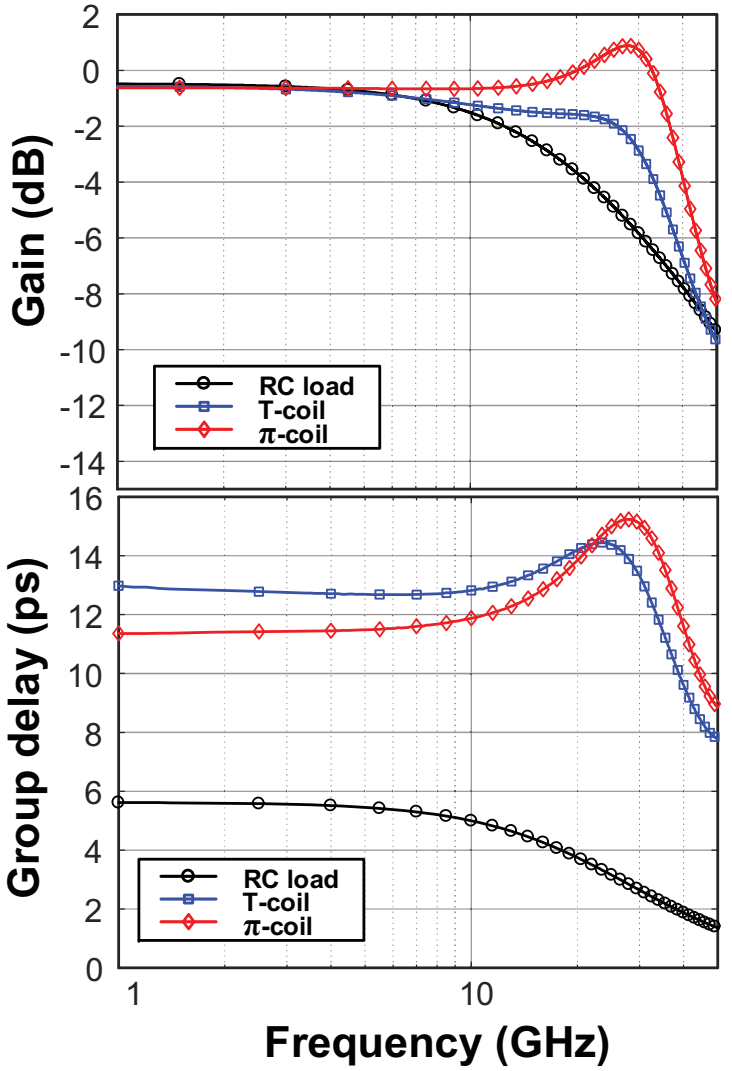
JTF – Jitter Transfer Function [25]

- TX model includes:
  - Finite precision FFE
  - DAC quantization & thermal noise
  - DAC nonlinearity
- Jitter to voltage noise conversion used to improve computational efficiency [20,21]
- Critical to include output/pad network

Electrical TX Model: Output Network



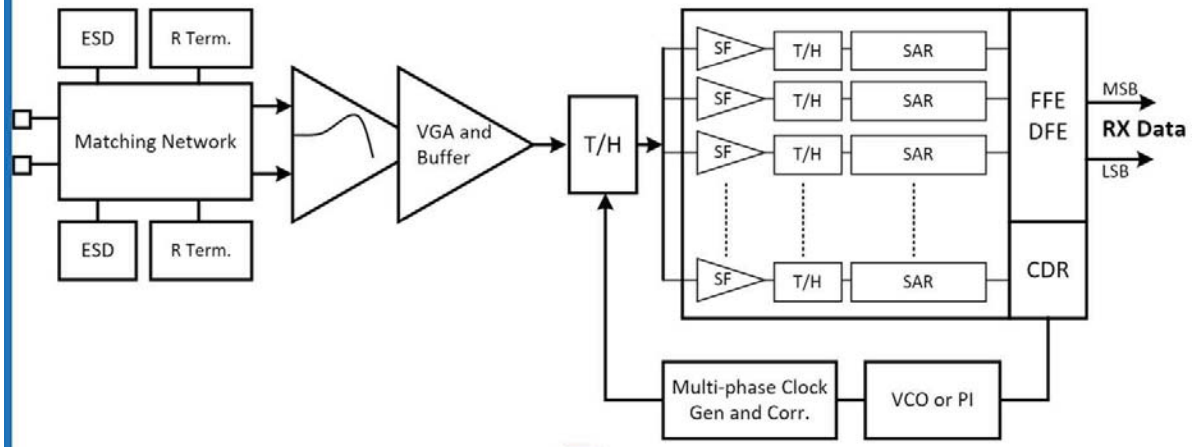
Ref. [26]



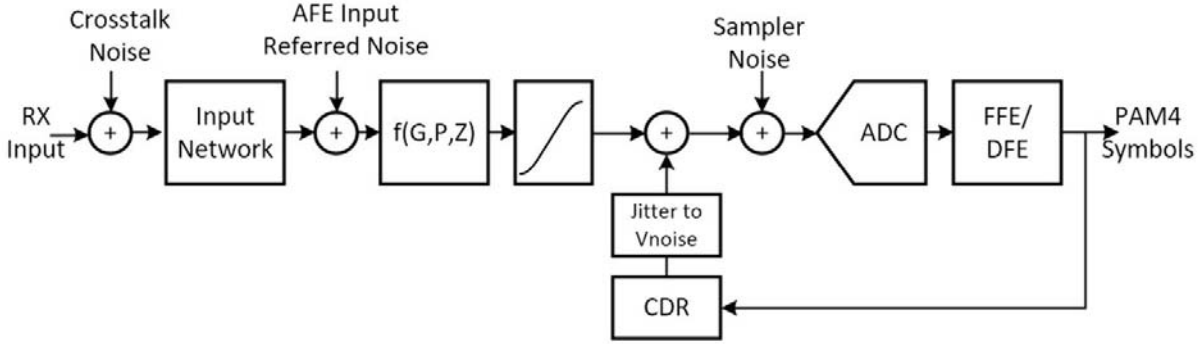
- Distributed output network to extend bandwidth in 56G/112G links
  - Group delay dispersion effects should be modeled

Electrical RX Model

Typical 112G RX block diagram

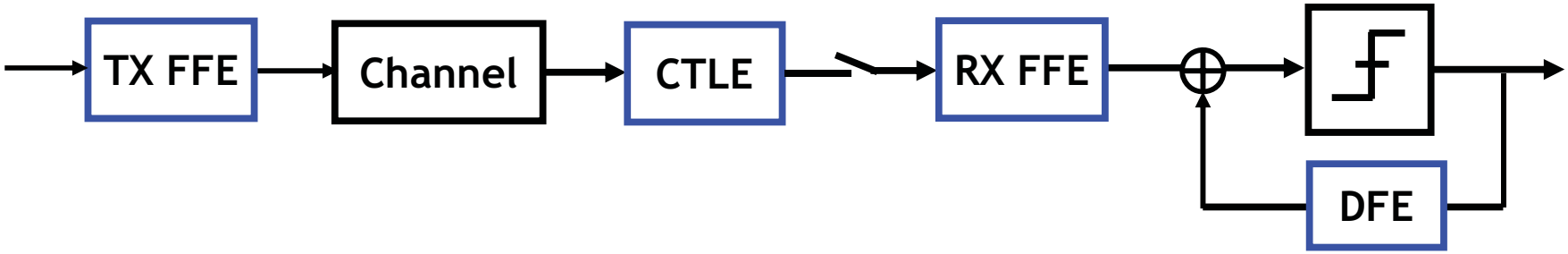


RX Model



- Analog front-end (AFE) model includes:
  - Pole-zero CTLE
  - Nonlinearity
  - Input-referred noise
- ADC = Ideal quantizer + input noise + clock jitter

Equalizer Optimization

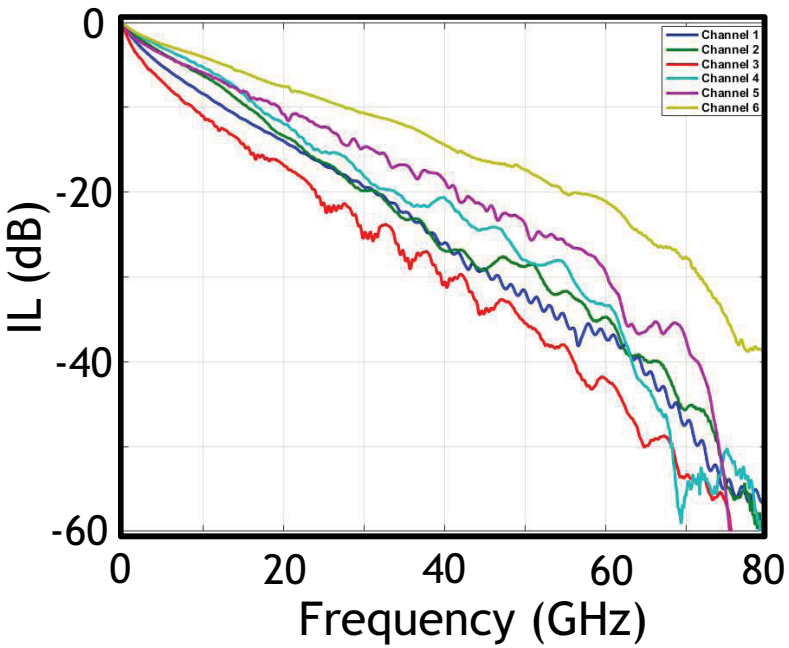


- Problem: Determine optimal TX FFE, CTLE, RX FFE+DFE for given channel and data rate
- Simplifications necessary to make problem tractable
- One approach:
  - Global search of CTLE with
    - TX FFE set to minimize MSE at ADC input
    - RX FFE minimizes ISI outside DFE span (in mean-square sense)
  - Choose setting with best FOM
    - E.g. SNR definition used in COM

Example 1: RX AFE Design

- Problem: Determine optimal analog front-end (AFE) design solution for a given set of target channels

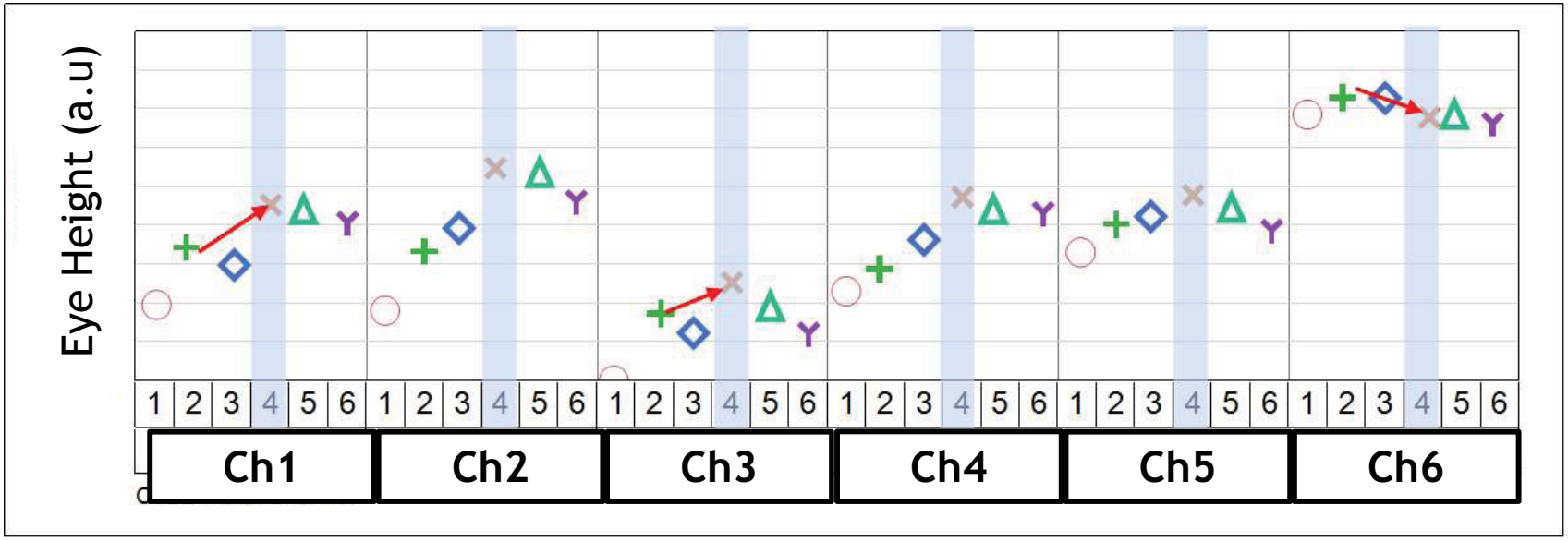
6 Target Channels



6 AFE options

AFE	CTLE Type	Post-CTLE Filter?
1	1z, 2p	No
2	1z, 2p	Yes
3	2z, 4p	No
4	2z, 4p	Yes
5	2z, (3rp + 1ccp)	Yes
6	2z, (2rp + 2ccp)	Yes

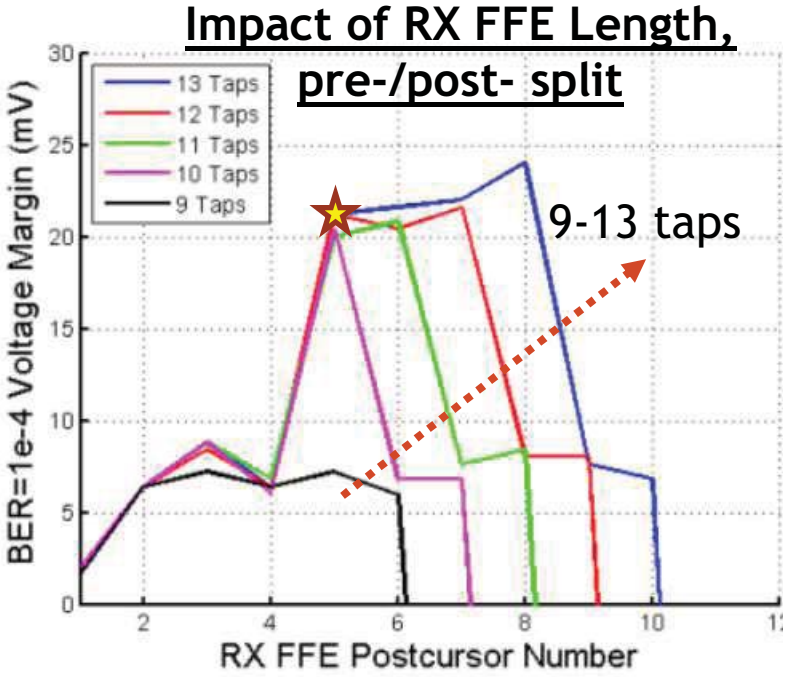
Example 1: AFE Design - Results

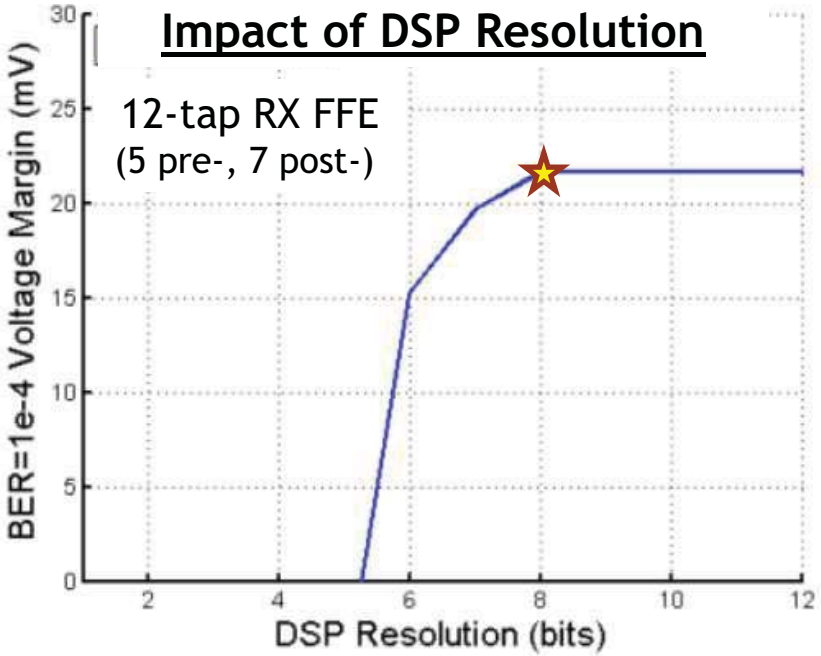


- Model results:
  - Addition of post-CTLE filter improves eye margins
  - Best margin with (2z, 4p) CTLE

AFE	CTLE Type	Filter?
1	1z, 2p	No
2	1z, 2p	Yes
3	2z, 4p	No
4	2z, 4p	Yes
5	2z, (3rp + 1ccp)	Yes
6	2z, (2rp + 2ccp)	Yes

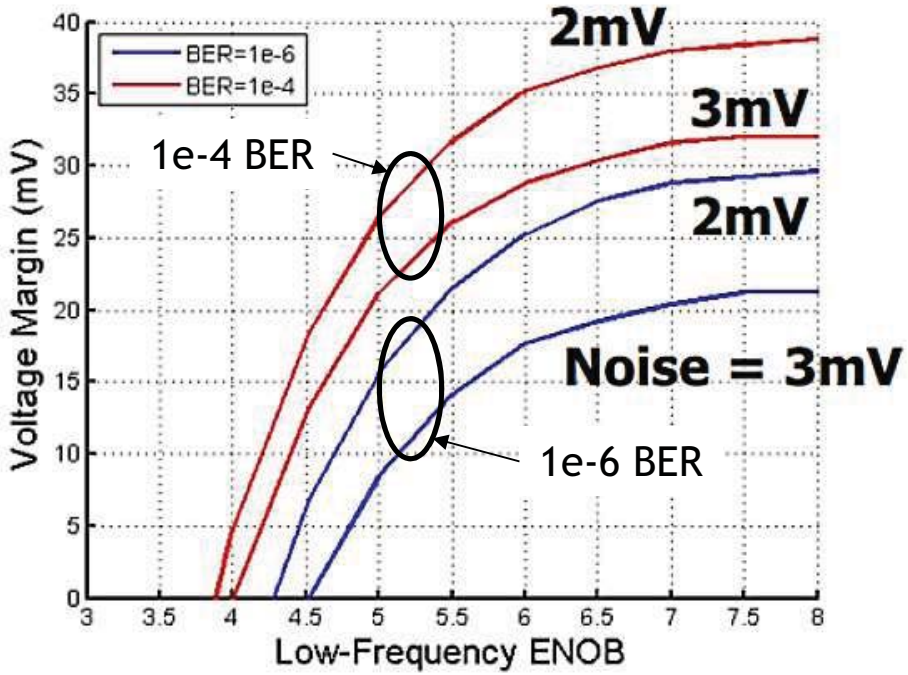
Example 2: ADC+DSP Sensitivity





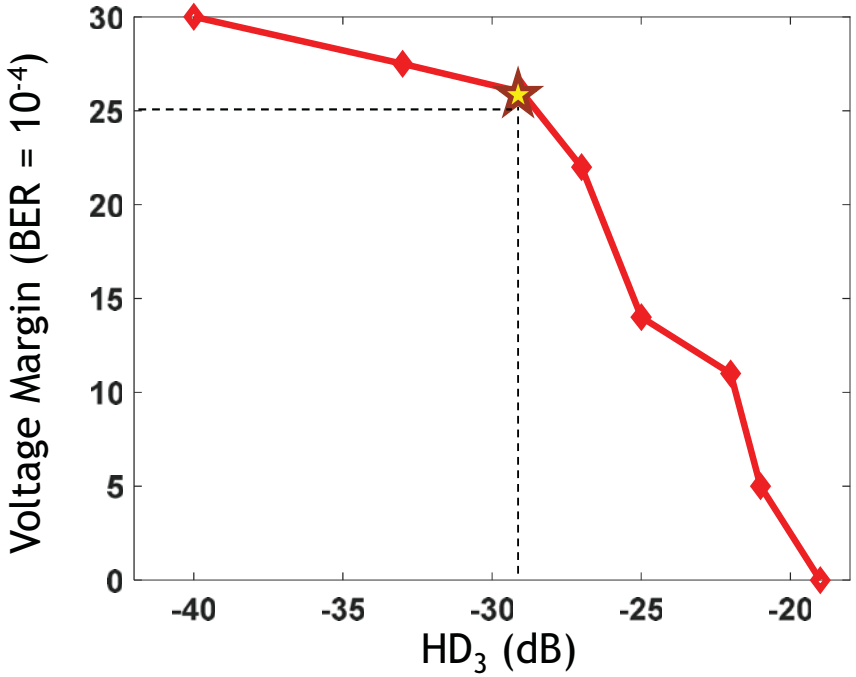
- Problem: Determine #FFE taps and resolution required in RX DSP for 112G SerDes
- Model results:
  - Approx. half of RX FFE taps should be post-cursor taps
  - DSP resolution = 8b

Example 2: ADC+DSP Sensitivity - Contd.



- Problem: Determine #FFE taps and resolution required in RX DSP for 112G SerDes
- Model results quantify ADC resolution – input noise tradeoff

Example 3: Sensitivity to AFE Linearity

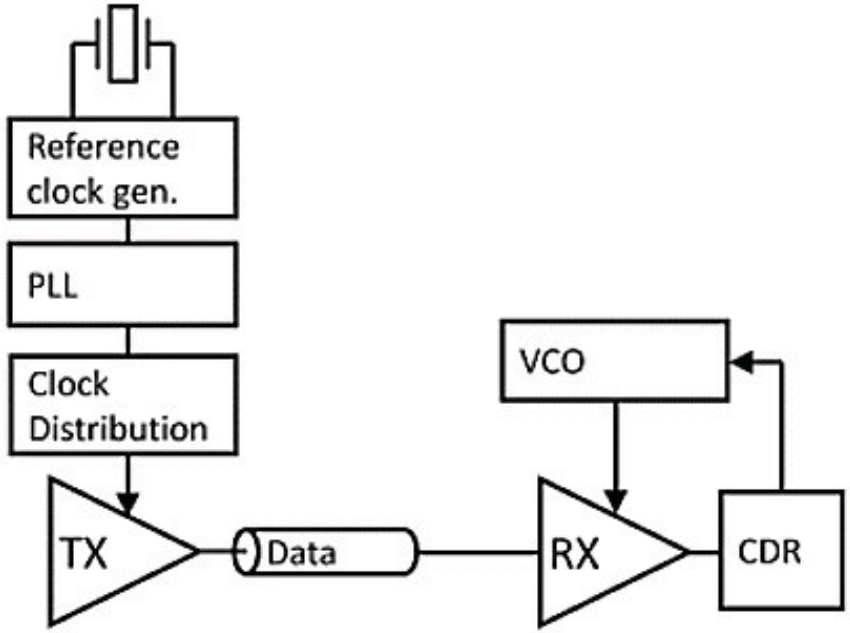


- Problem: Determine AFE linearity requirement based on link-level specifications
- Results show HD3 < ~-30 dB sufficient to ensure 25 mV margin:
  - Can help optimize transceiver power by avoiding AFE over-design

Outline

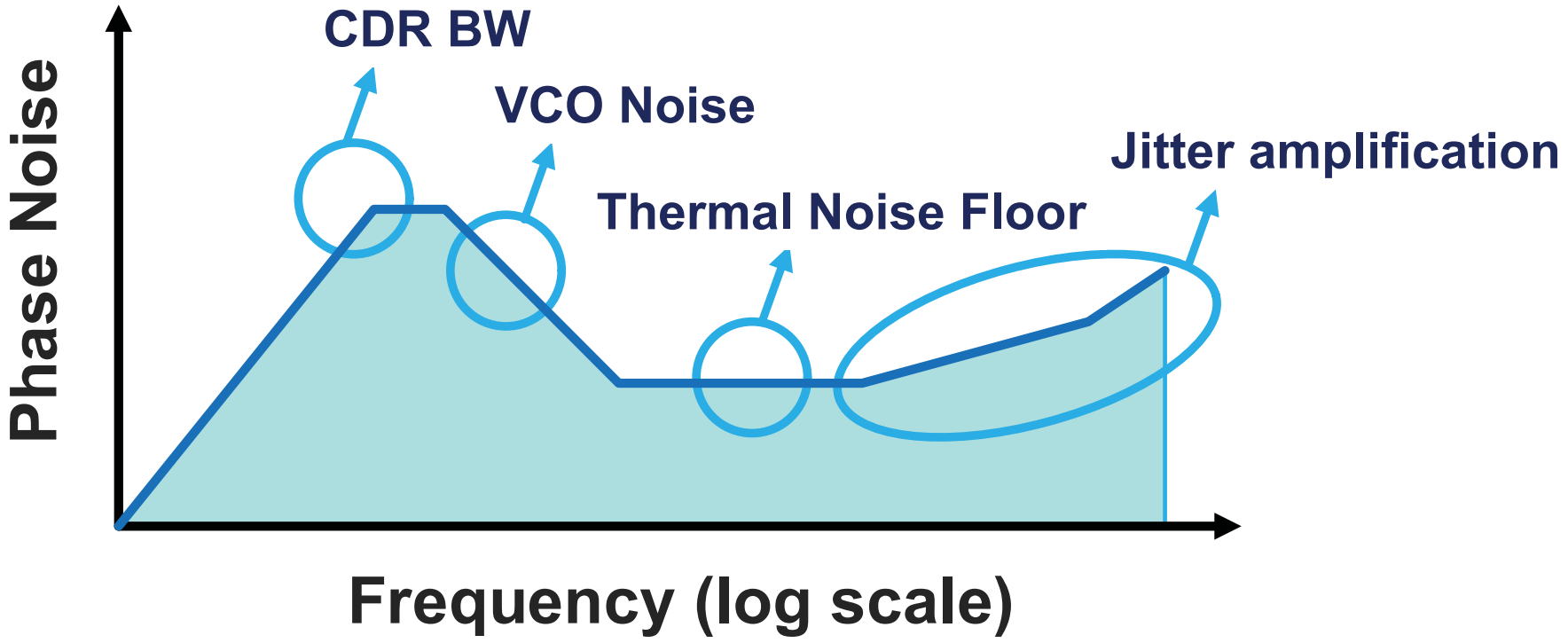
- Introduction
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
  - Modeling Methods
  - Electrical Links
  - Clocking
  - Optical Links
- Conclusion

Clocking Analysis



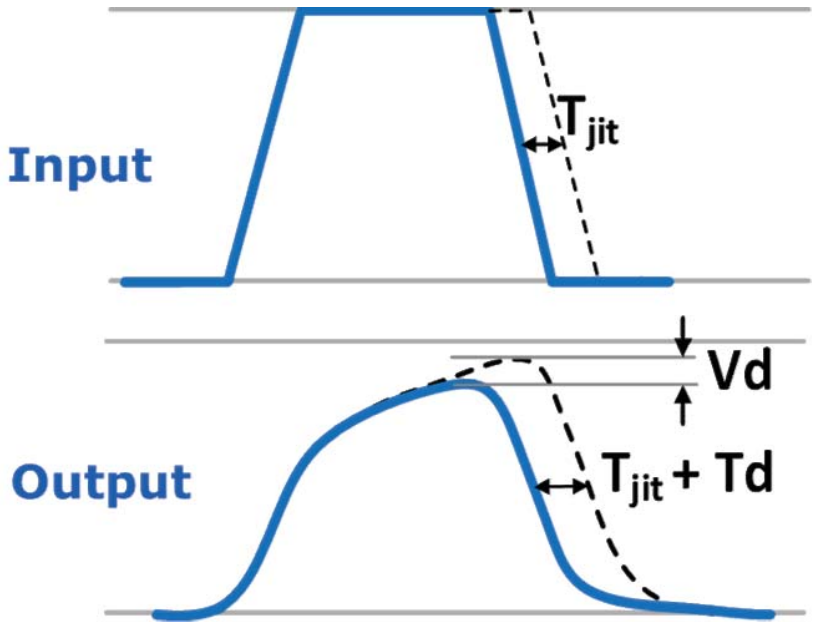
- Used to design TX, RX clocking solutions to meet jitter specifications [25]
  - E.g. PLL bandwidth, clock fan-out, #stages etc..
- Random jitter optimization is critical for 56G/112G links

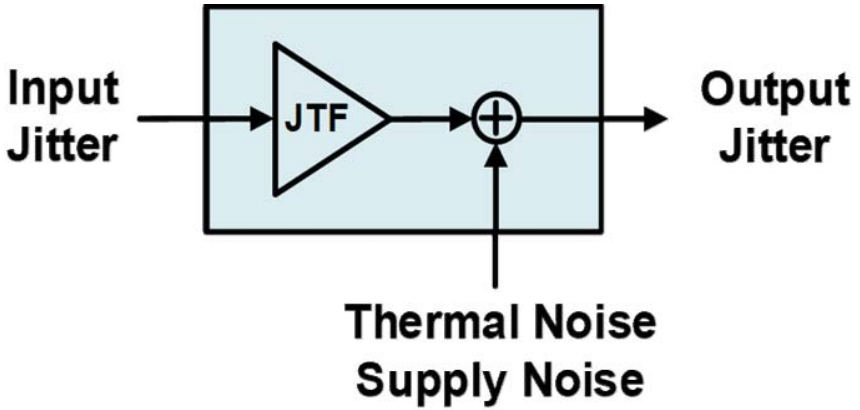
Random Jitter Frequency Profile



- Increasing clock frequency → Increasing RJ from clock *distribution*

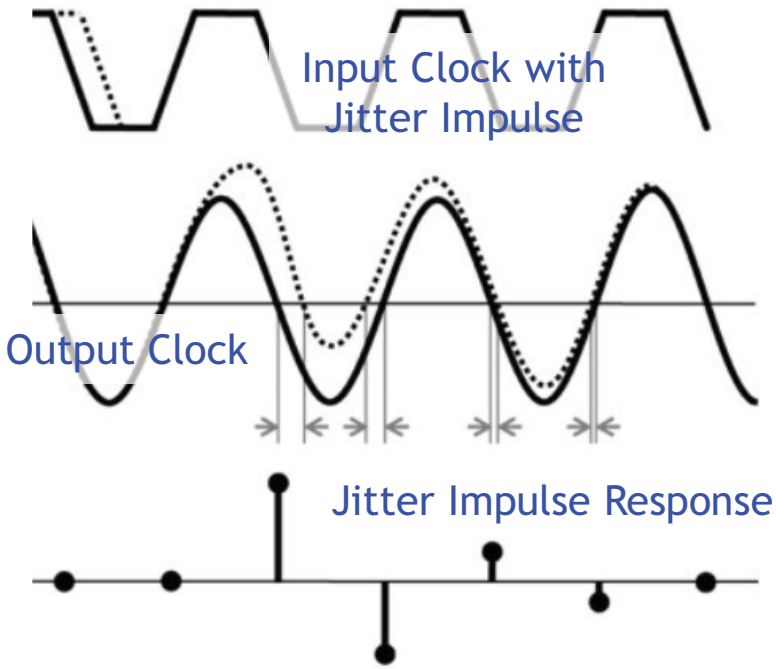
Jitter Amplification

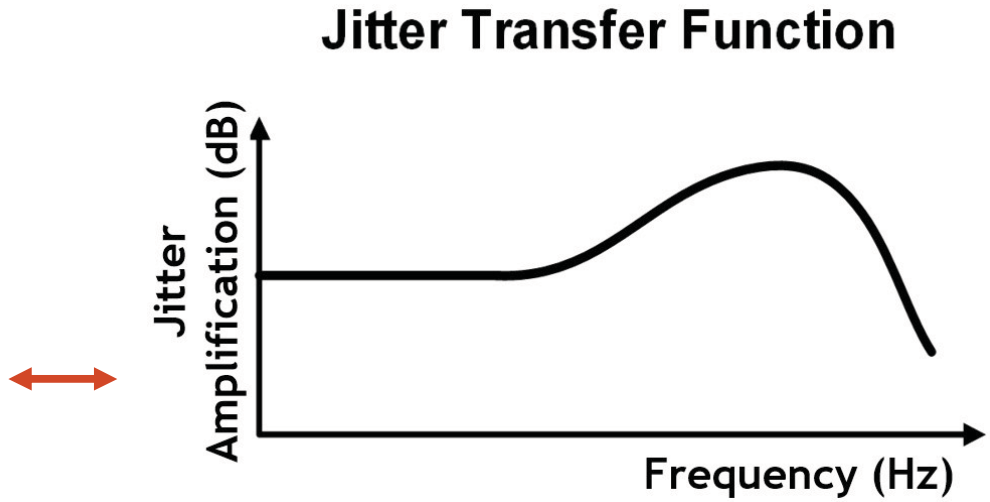




- Bandwidth-limited circuits amplify high frequency jitter
- This can be analyzed using the Jitter Transfer Function (JTF)

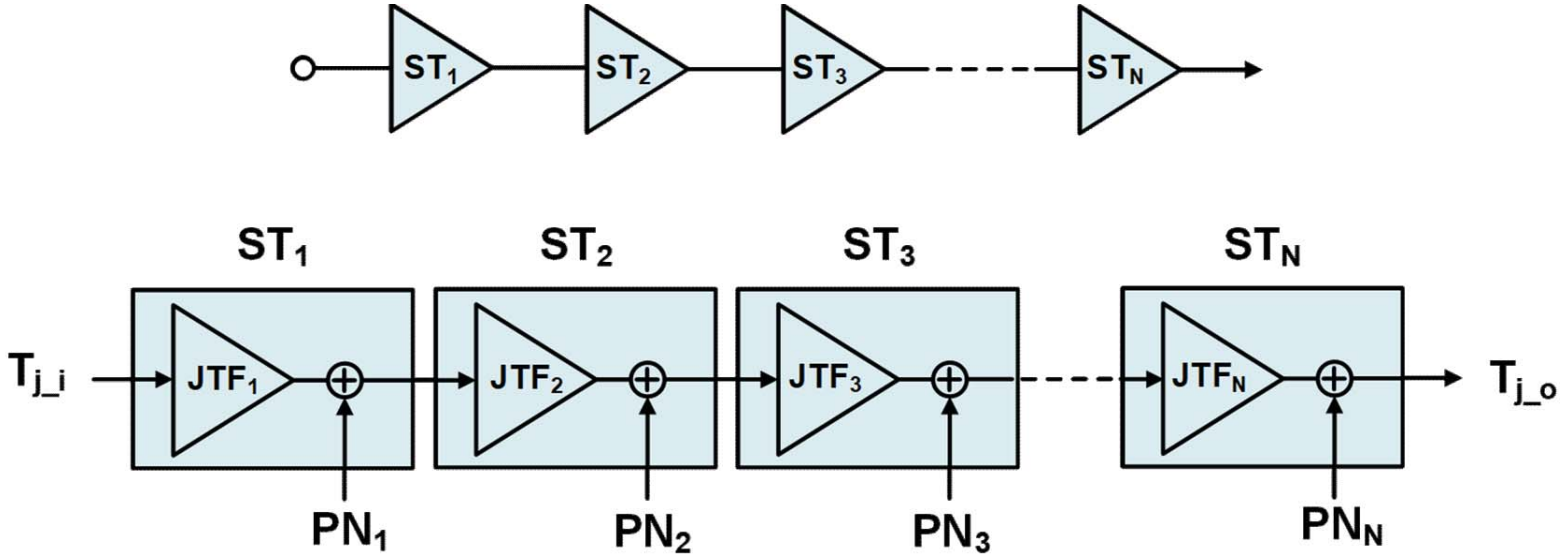
Jitter Transfer Function





- Capture jitter impulse response (JIR) by shifting one clock edge
- Z-transform of JIR = Jitter Transfer Function

Clock Distribution Analysis



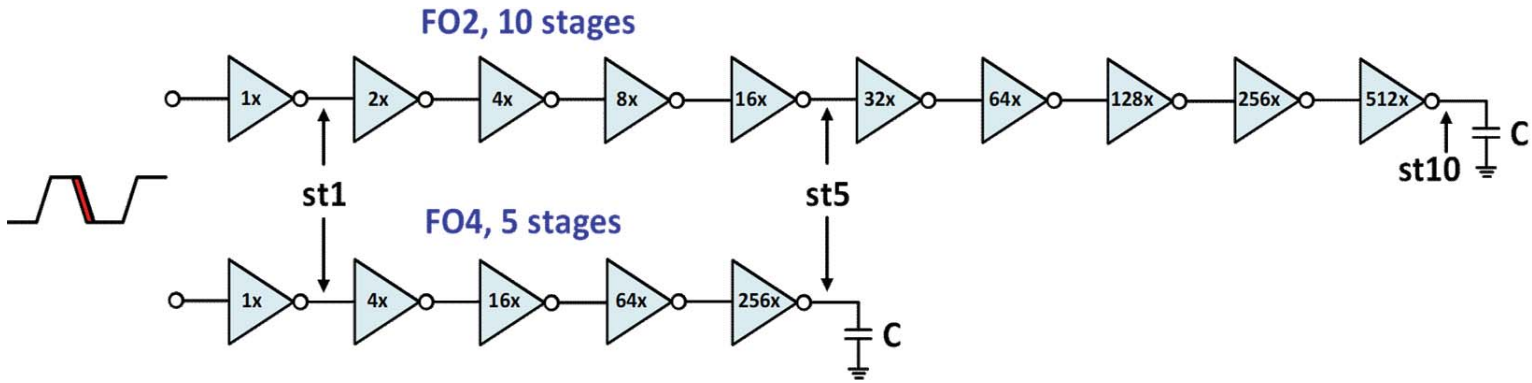
PN_i = Phase noise of i-th stage

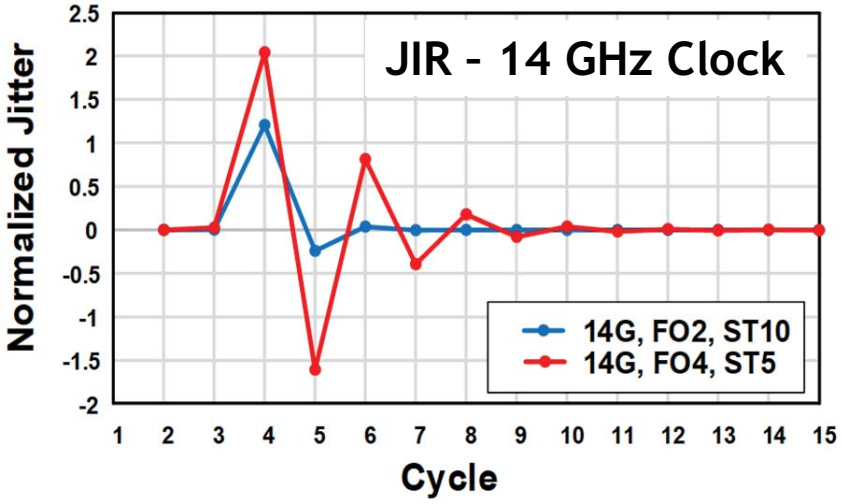
$$T_{j\_o}=(\ldots (((T_{j\_i} * JTF_1) + PN_1) * JTF_2) + PN_2) * \ldots) * JTF_N) + PN_N$$

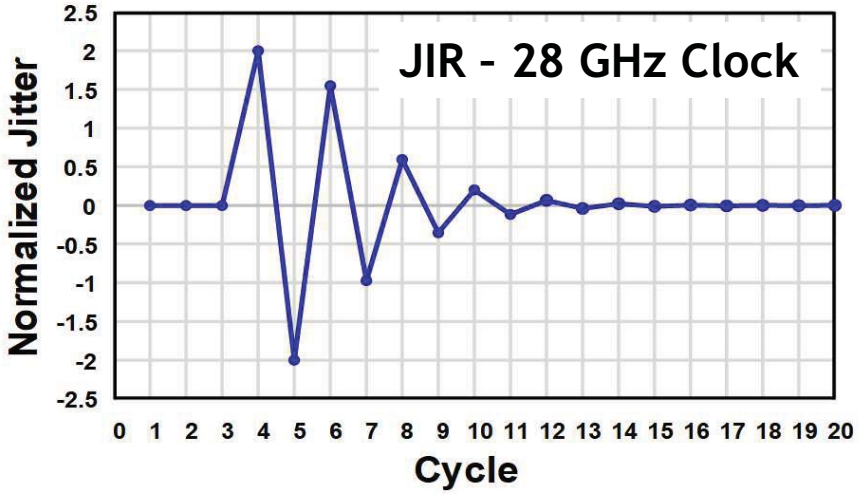
- Jitter transfer functions enable efficient analysis of clock distribution options

Clocking Analysis Example

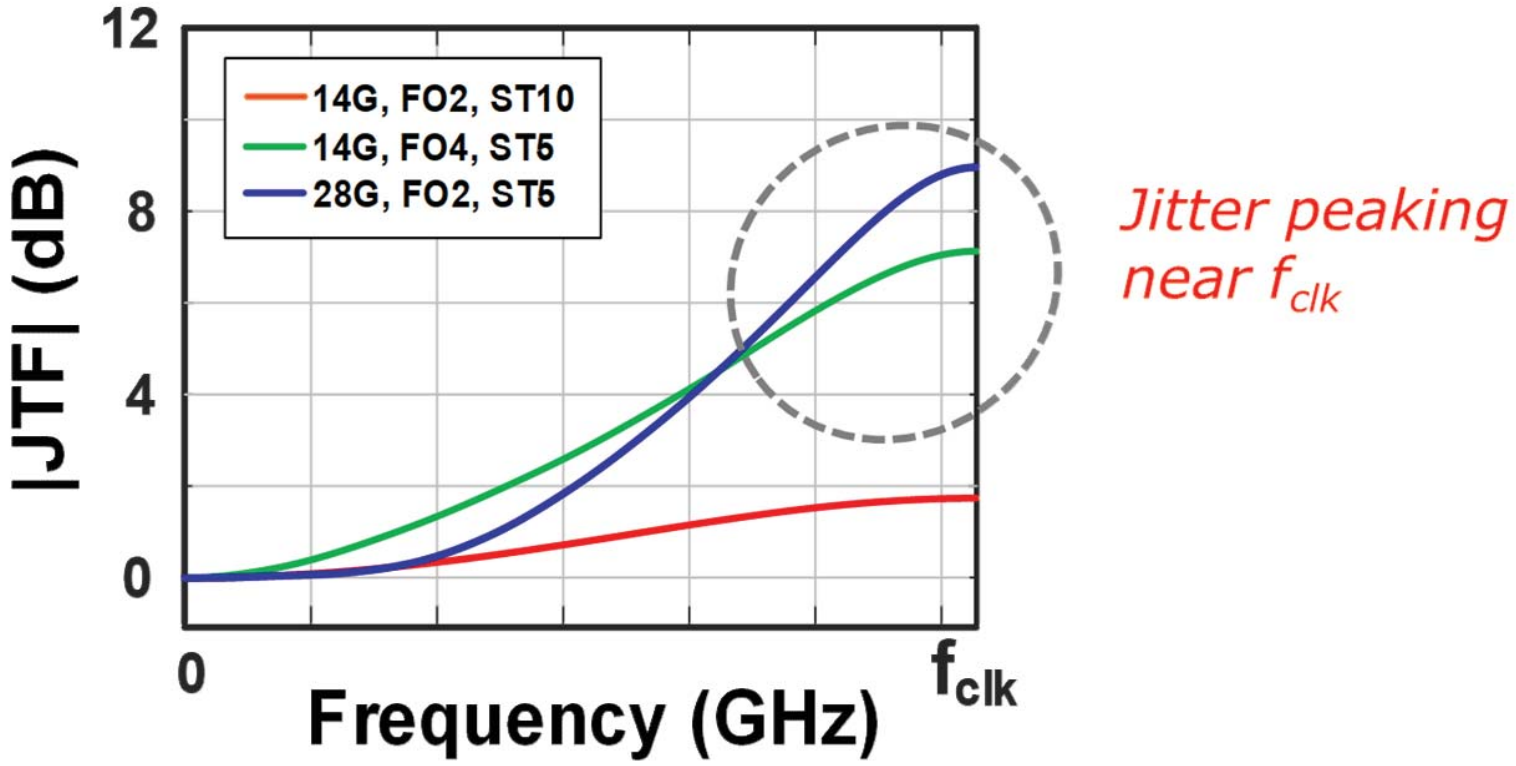
- Compare 2 clock distribution scenarios:







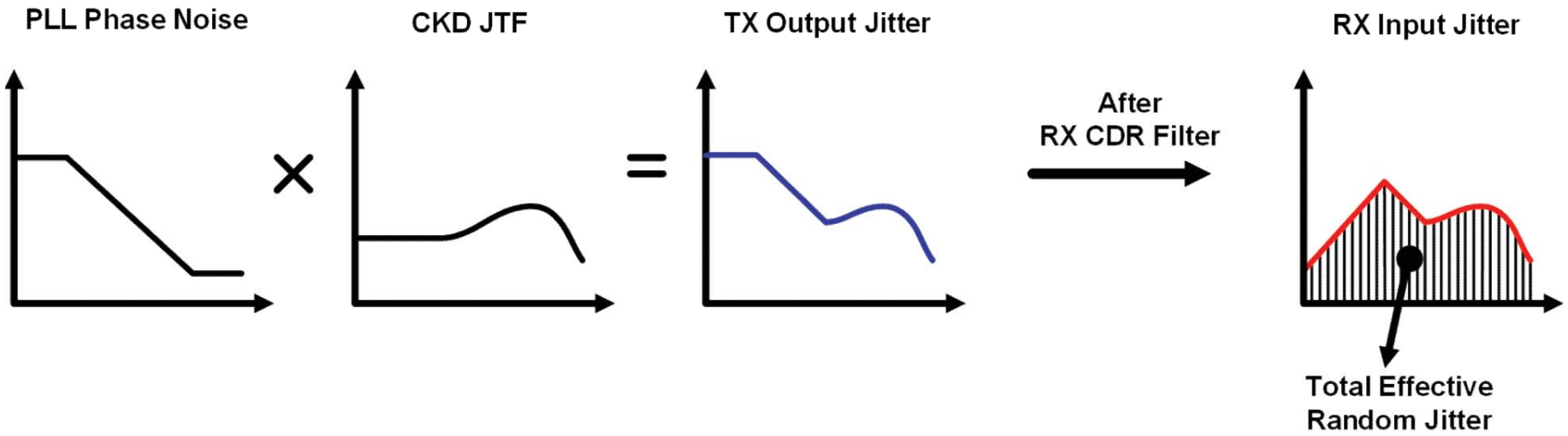
Clocking Analysis Example - Contd.



- JTF analysis can guide clocking scheme selection and clock distribution design

TX Total Jitter Computation

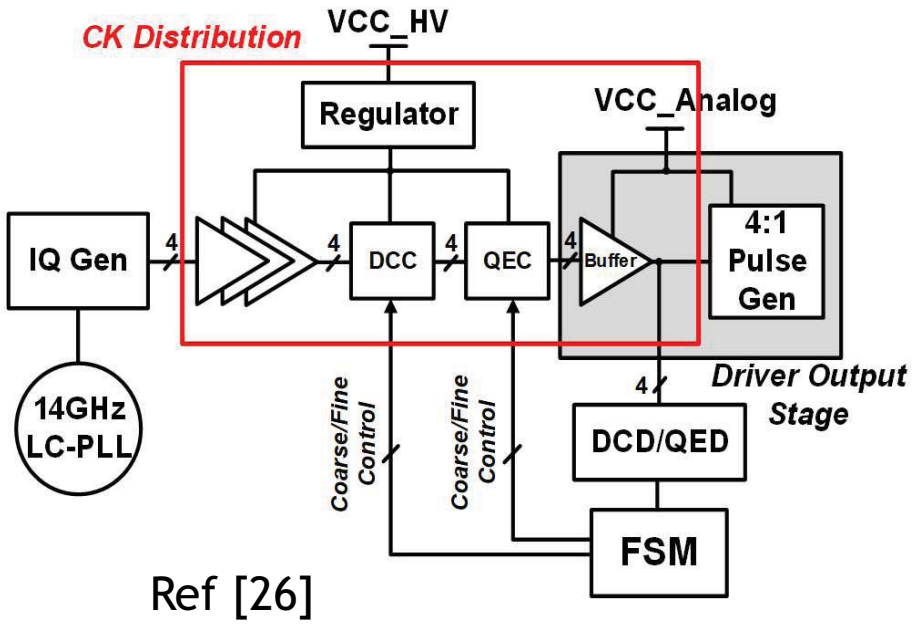
- TX jitter frequency spectrum can be computed from:
  - PLL phase noise profile
  - Clock distribution JTF, phase noise
  - CDR filter BW



- Dominant jitter sources can be quickly identified using this method.

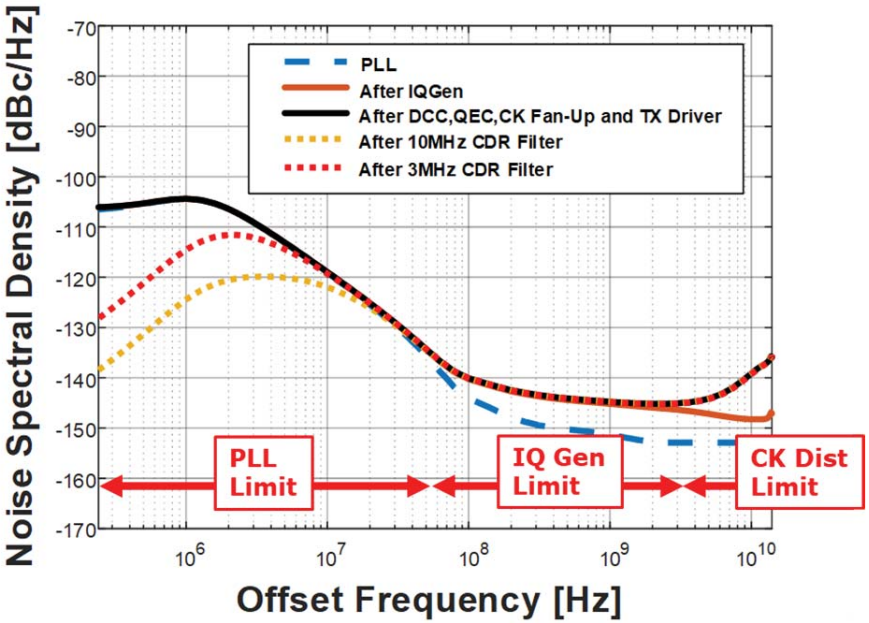
Design Example: 112G TX Clocking

Clocking for 112G TX



Ref [26]

Simulated Jitter PSD



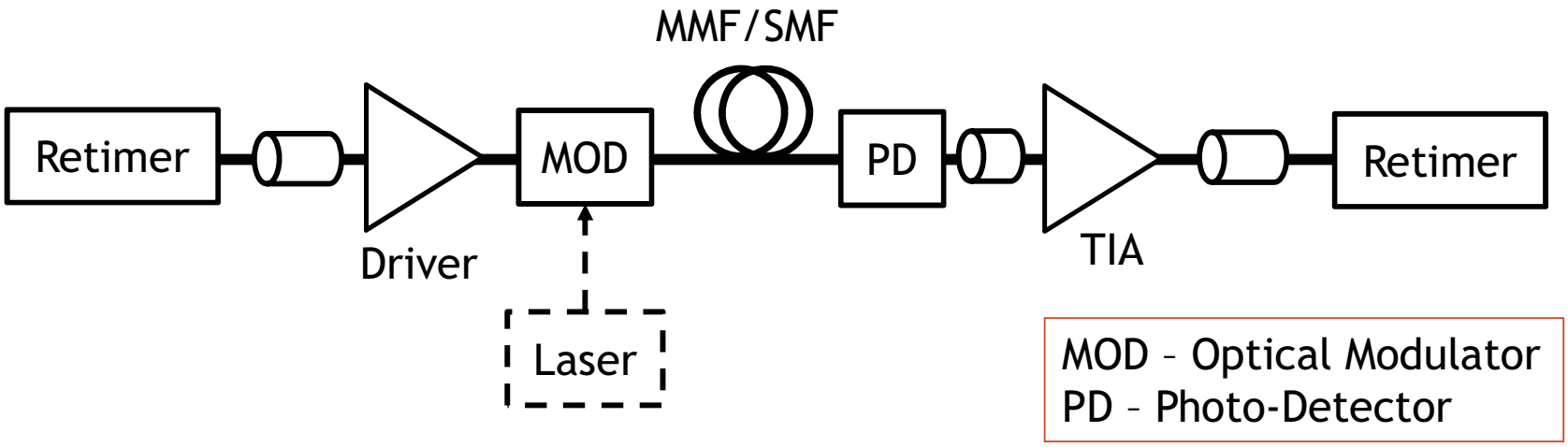
RJ after 10 MHz CDR = 185 fs-rms

- RJ computed from full clock path simulation matches that from JTF analysis

Outline

- Introduction
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
  - Modeling Methods
  - Electrical Links
  - Clocking
  - Optical Links
- Conclusion

Optical Link Modeling

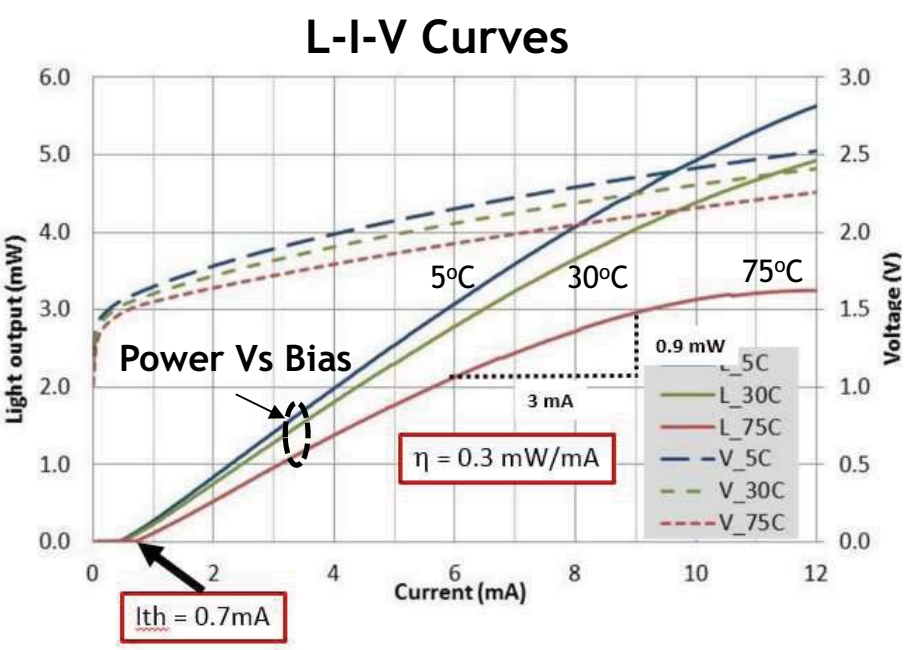


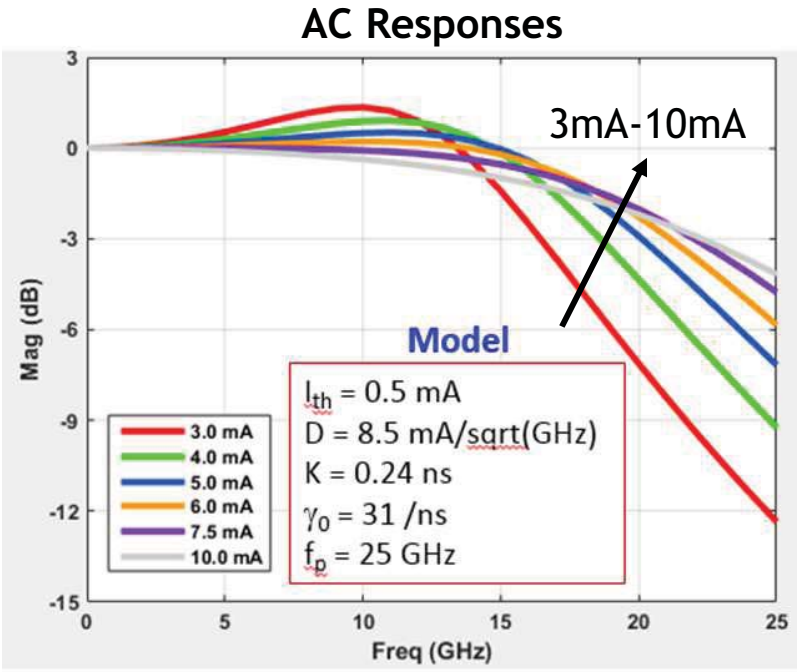
- Includes several components from electrical links
- Key new components:
  - Optical modulator
  - Photodetector (PD)
- Several commercial tools now available [27, 28]

Optical Modulator Modeling

- Two types of intensity modulators:
  - Direct modulators which modulate laser directly
    - E.g. VCSEL
  - External modulators which modulate CW laser output
    - E.g. Mach-Zehnder Modulator (MZM), Electro-Absorption Modulator (EAM), Ring Modulator
- Key parameters:
  - Bandwidth
  - Linearity
  - Noise (RIN)

Example Modulator Model: VCSEL





Nonlinear VCSEL Model [29]

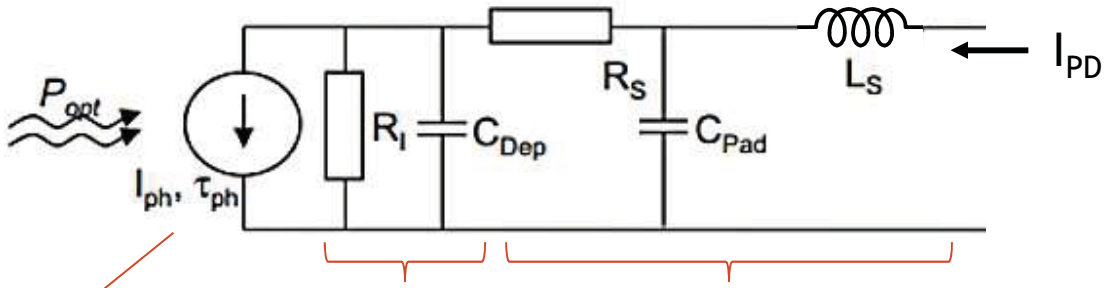
$$H(f) = const \times f_r^2/(f_r^2 - f^2 + j(f/2\pi)\gamma)$$

$$f_r = D\sqrt{I_{VCSEL} - I_{th}}$$

$$\gamma = Kf_r^2 + \gamma_o$$

- Bias-dependent transfer function
  - Parameters extracted from L-I-V and S-parameter data
  - Results in asymmetric PAM4 eye

Photodetector Model



$$I_{ph} = P_{opt} \cdot R \cdot \frac{1}{1 + j \cdot \omega \cdot \tau_{ph}}$$

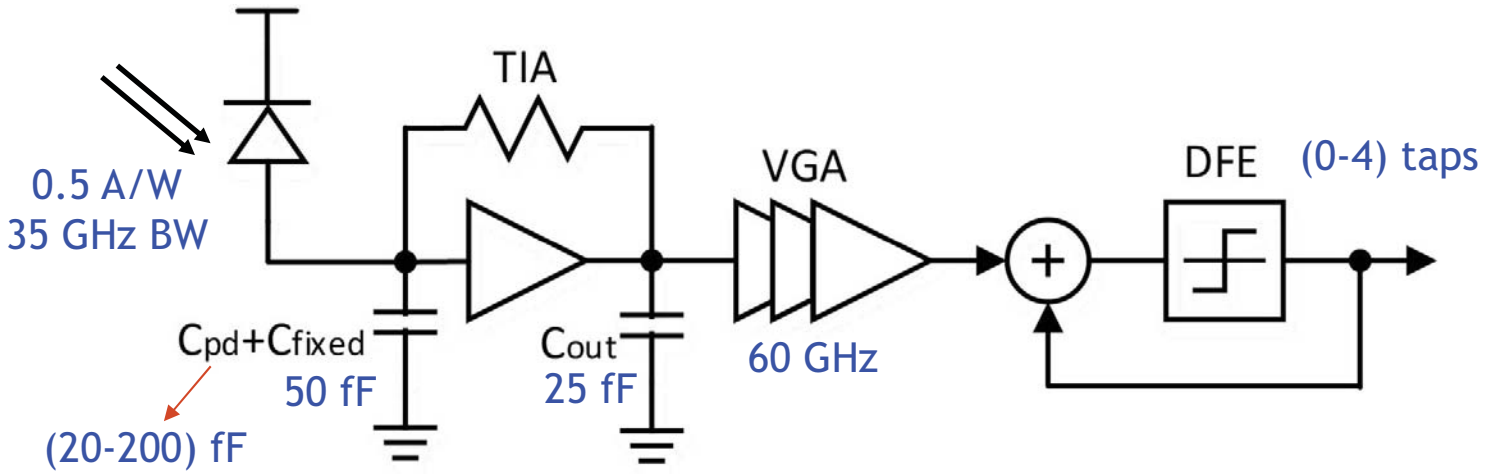
Transit time-limited bandwidth

P_{opt} - Incident optical power
R - Responsivity
τ_{ph} - Photon lifetime

- Key parameters and typical values:
  - Responsivity: 0.5-0.8 A/W
  - Bandwidth: >35 GHz for 50 GBaud links
  - Capacitance: <70 fF

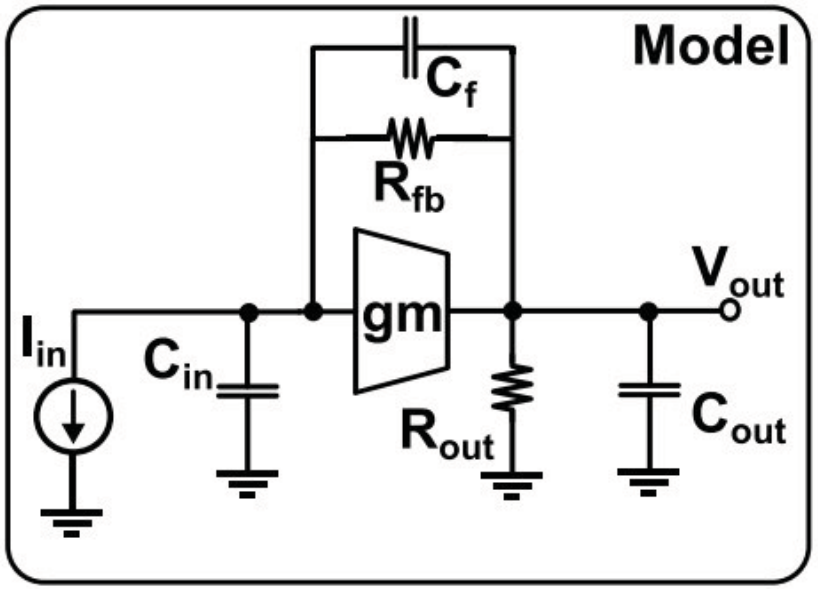
Example: TIA Design

56 GBd Optical Receiver (Target BER = 5e-5)



- Goal: Determine TIA specifications to optimize receiver sensitivity
  - Need to co-optimize TIA and DFE
  - Also comprehend sensitivity to PD capacitance (C_{pd})

Example: TIA Design (Modeling)



Ref [30]

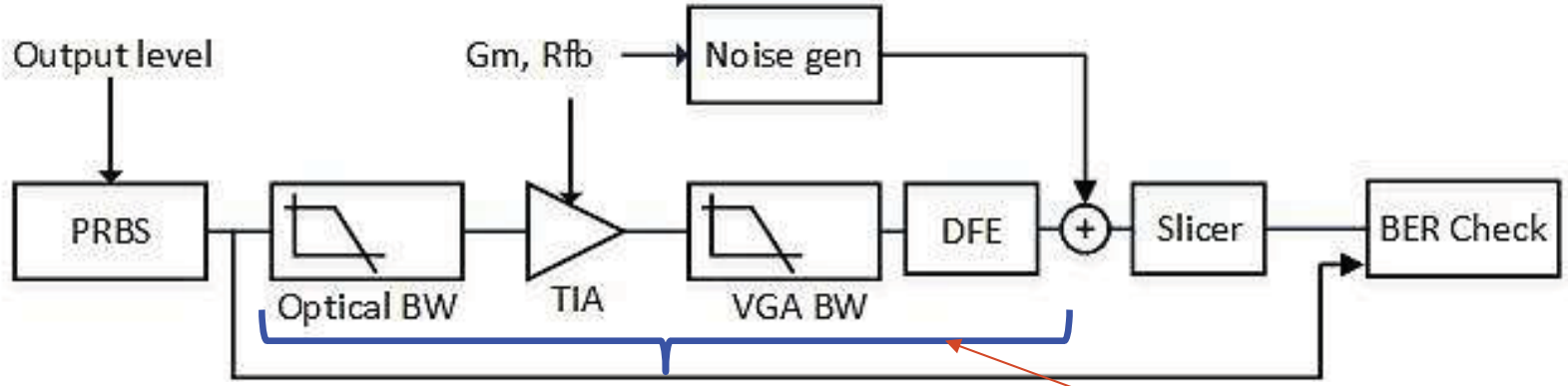
$$Z_t(s) = \frac{R_{\text{out}}(gmR_{\text{fb}} - 1 - sC_fR_{\text{fb}})}{1 + gmR_{\text{out}} + sD_{t1} + s^2D_{t2}}$$

$$D_{t1} = C_{\text{in}}(R_{\text{out}} + R_{\text{fb}}) + C_fR_{\text{fb}}(1 + gmR_{\text{out}})$$
$$D_{t2} = R_{\text{fb}}R_{\text{out}}C_{\text{in}}(C_{\text{out}} + C_f).$$

$$A = gmR_{\text{out}}$$
$$f_t = \frac{gm}{2\pi C_{\text{gate}}}$$

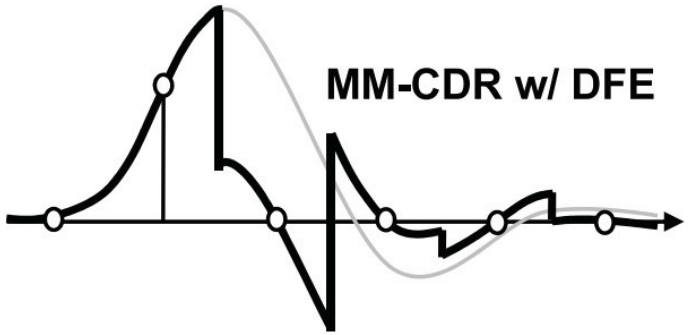
- Shunt-feedback TIA behavior can be described analytically
  - Simulated step response can be used for more complex TIAs

Example: TIA Design (Analysis)



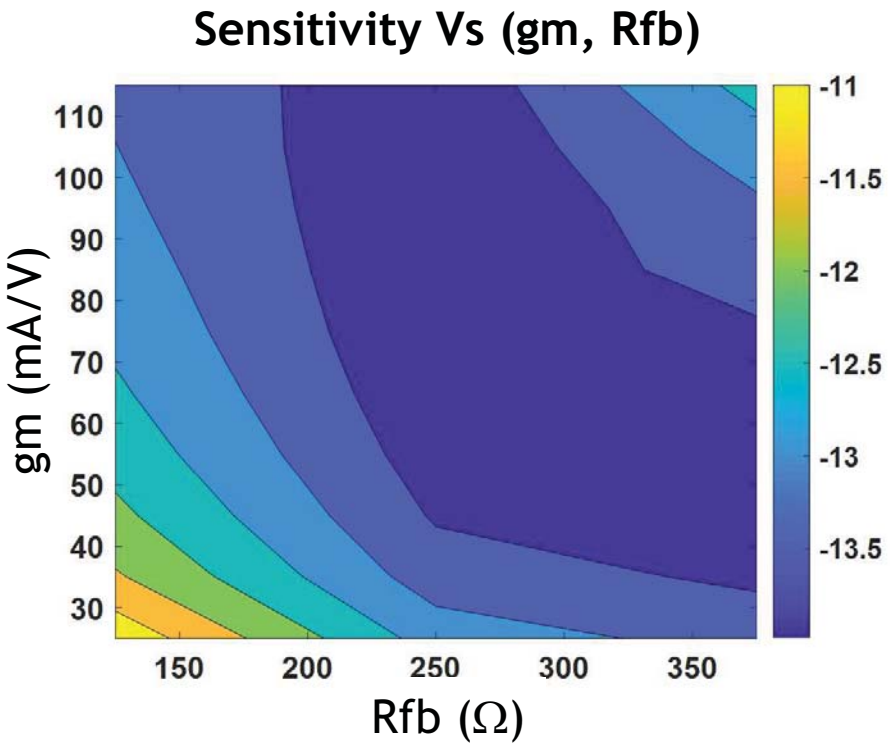
Combined pulse response can be analytically computed

- For each design choice (g_m, R_{fb})
  - Compute overall PR
  - Find sampling point from CDR timing function
  - Create waveform at slicer input by convolution
  - Find RX sensitivity
    - Decrease PRBS level until BER>target

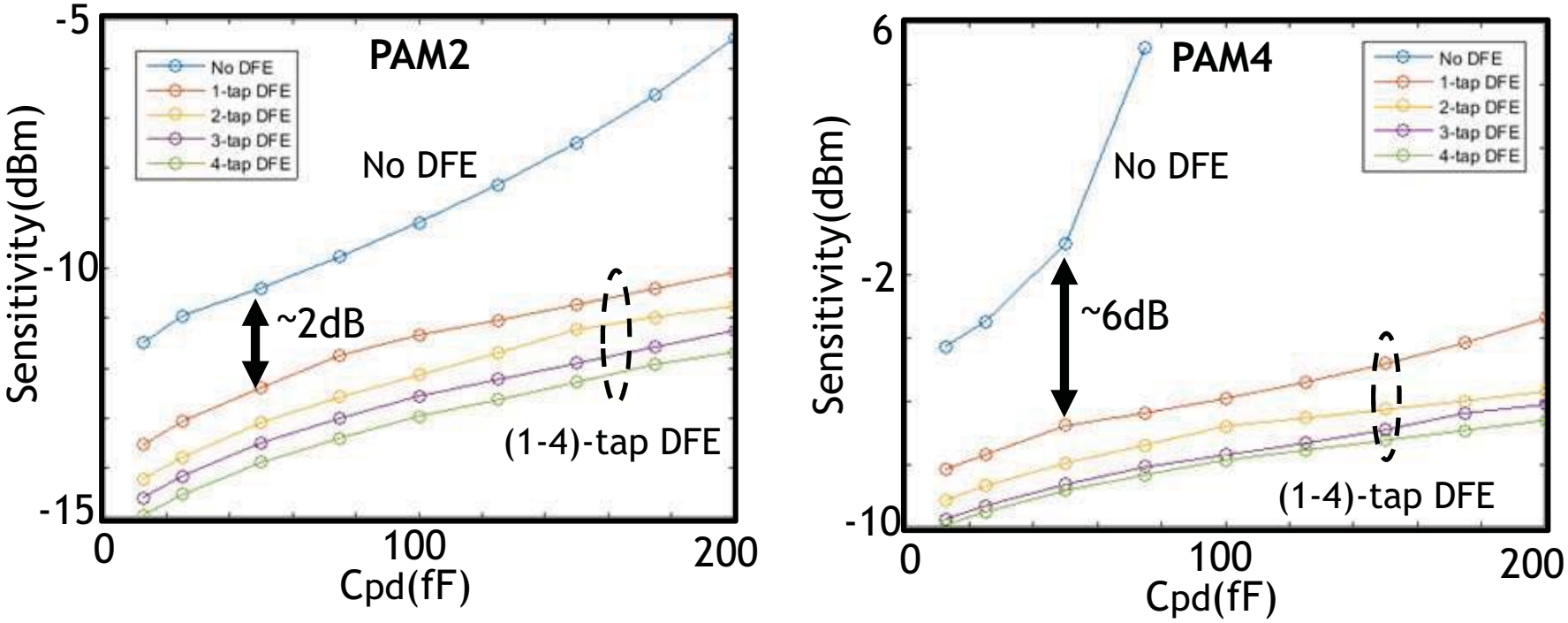


Example: TIA Design (Analysis - Contd.)

- Optimized design point search: For each DFE length and PD capacitance,
  - 2-D sweep of Gm and Rfb to find optimal sensitivity



Example: TIA Design (Results)



- DFE improves RX sensitivity by >2 dB
  - And reduces sensitivity to PD cap
- DFE benefit is more significant for PAM4

Outline

- Introduction
- 56G,112G Standards
- Link Budgeting
- Link Modeling & Analysis
- Conclusion

Conclusion

- Data center bandwidth needs driving demand for 56G/112G electrical and optical links
- 2 key standards: OIF-CEI & IEEE 802.3 (Ethernet)
  - NRZ to PAM4 transition has led to new specifications
  - Most 56G standards complete. Several 112G standards under development.
- Channel Operating Margin (COM) is a standardized technique for link budgeting
- Link models enable efficient design exploration and system level optimization
  - Reviewed several examples using time-domain models

Acknowledgements

- The authors thank the following contributors for their assistance with this work:
  - Frank O'Mahony
  - Shiva Kiran
  - Jihwan Kim
  - Bryan Casper
  - Hai-Feng Liu

  } Intel

  - Prof. Samuel Palermo (Texas A&M University)

References - I

1. Cisco Visual Networking Index: Forecast and Trends, 2017-2022 White Paper, https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html
2. A. Singh et al., “Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google’s Datacenter Network”, Sigcomm 2015
3. Optical Internetworking Forum (OIF), https://www.oiforum.com/
4. IEEE 802.3 Ethernet Working Group, http://www.ieee802.org/3/
5. OIF-CEI-04.0, “Common Electrical I/O (CEI) - Electrical and Jitter Interoperability agreements for 6G+ bps, 11G+ bps, 25G+ bps I/O and 56G+ bps”, Dec. 2017.
6. IEEE Standard 802.3bs, Amendment 10: Media Access Control Parameters, Physical Layers, and Management Parameters for 200 Gb/s and 400 Gb/s Operation, Dec. 2017.
7. Ethernet Alliance, https://ethernetalliance.org/
8. H. Zhang et al., “PAM4 Signaling for 56G Serial Link Applications – A Tutorial”, DesignCon 2016.
9. M. Pisati et al., “A Sub-250mW 1-to-56Gb/s Continuous-Range PAM-4 42.5dB IL ADC/DAC-Based Transceiver in 7nm FinFET”, ISSCC 2019.
10. J. King, “TDECQ and SECQ vs Rx sensitivity: review of previous presentations and proposed changes,” IEEE Interim, P802.3cd, Jan 2018
11. H. Li et al.,”A 112 Gb/s PAM4 Transmitter with Silicon Photonics Microring Modulator and CMOS Driver,” Th4A.4 OFC 2019.
12. S. Bhoja et al., “Precoding proposal for PAM4 modulation,” IEEE 802.3 100 Gb/s Backplane and Cable Task Force, Sept. 2011.

References - II

13. G. Zhang et al., “Impact of DFE Error Propagation and Precoding on FEC Performance for PAM4 Link Systems,” ICITEE 2018.
14. R. Mellitz, “Channel Operating Margin Tutorial,” http://www.ieee802.org/3/cb/public/mar16/mellitz_3cb_01_0316.pdf
15. Ref: “Measuring Channel Operating Margin,” Anritsu White Paper.
16. M. Li et al., “112 Gbps LR COM Updates,” OIF2018.417.00, Nov 2018.
17. B. Casper et al., “An accurate and efficient analysis method for multi-Gb/s chip-to-chip signaling schemes ”, VLSI Circuits Symposium, 2002.
18. B. Casper, et al., “Future microprocessor interface: Design, analysis and optimization”, CICC 2007
19. A. Sanders, M. Resso, and J. D’Ambrosia, “Channel Compliance Testing Utilizing Novel Statistical Eye Methodology,” DesignCon 2004
20. V. Stojanovic et al., “Modeling and analysis of high speed links ”, CICC 2003
21. G. Balamurugan et al., “Modeling and Analysis of High-Speed I/O Links”, IEEE Transactions on Advanced Packaging, Apr 2009.
22. M. Li, et al., “Advancements in High-Speed Link Modeling and Simulation,” IEEE Custom Integrated Circuits Conf., 2013
23. M.-J. Park, et al., "Fast and Accurate Event-Driven Simulation of Mixed-Signal Systems with Data Supplementation," IEEE Custom Integrated Circuits Conference (CICC), Sept. 2011.

References - III

24. IBIS-AMI standard website: http://www.vhdl.org/ibis
25. B. Casper et al., "Clocking analysis, implementation and measurement techniques for high-speed data links - a tutorial," in IEEE Transactions on Circuits and Systems (T-CAS), Vol 56, No. 1, Jan 2009, pp.17-39.
26. J. Kim et al., "A 112Gb/s PAM-4 transmitter with 3-tap FFE in 10nm CMOS," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2018, pp. 102-103.
27. Lumerical INTERCONNECT, https://www.lumerical.com/products/interconnect/
28. RSoft, https://www.synopsys.com/optical-solutions/rsoft.html
29. M. Raj et al., “A Modelling and Nonlinear Equalization Technique for a 20 Gb/s 0.77 pJ/b VCSEL Transmitter in 32 nm SOI CMOS”, IEEE JSSC, July 2016.
30. I. Ozkaya et al., “A 64-Gb/s 1.4-pJ/b NRZ optical receiver data-path in 14-nm CMOS FinFet,” IEEE J. Soild-State Circuits, vol. 52 , no. 12, pp. Dec. 2017.