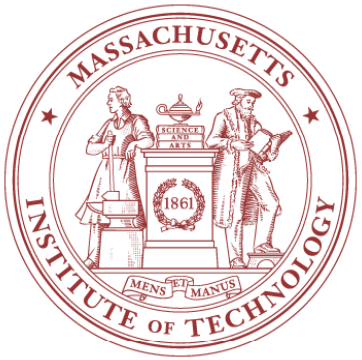
High-speed serial links: Design Trends and Challenges

Vladimir Stojanović



Integrated Systems Group

Massachusetts Institute of Technology

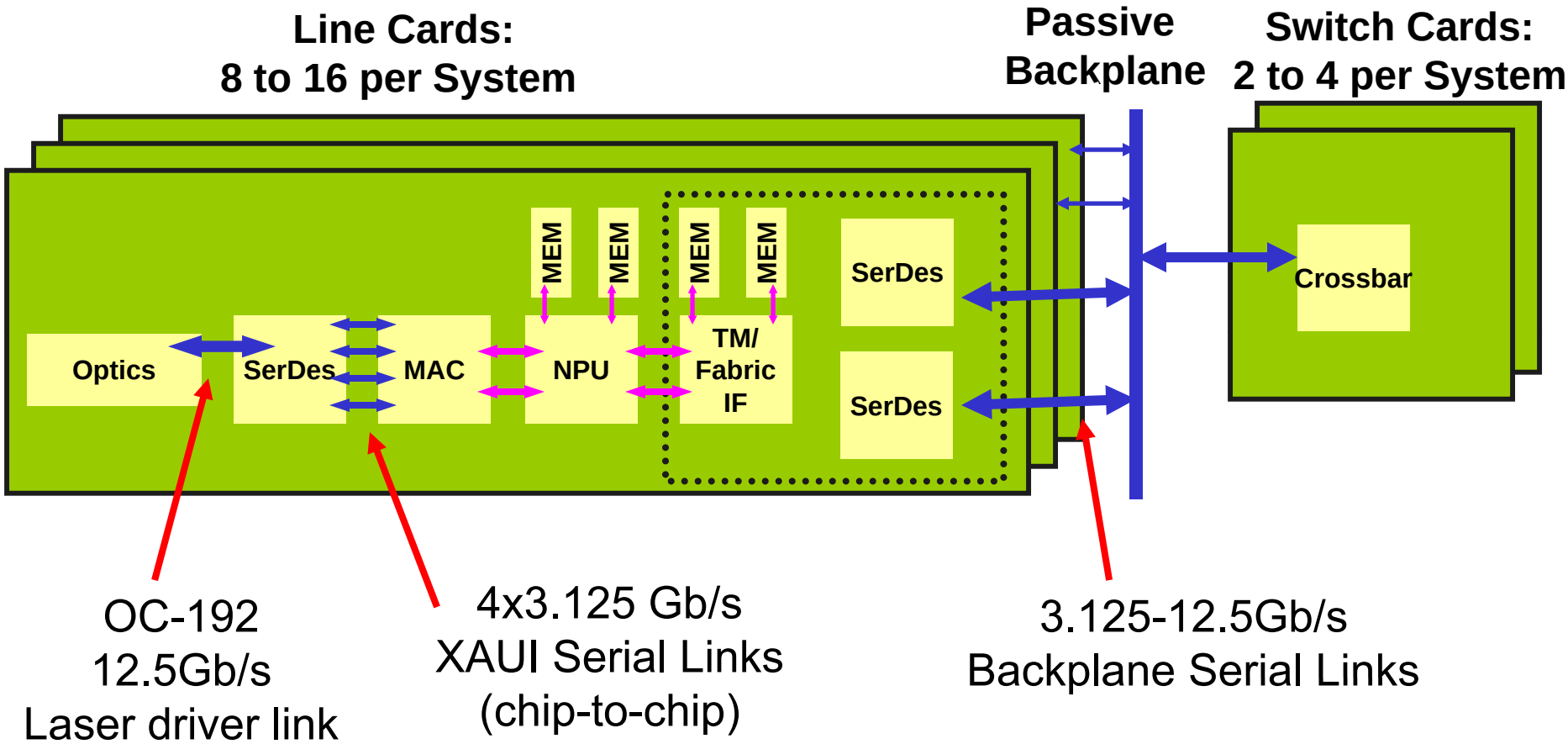
Backbone router – lots of high-speed links

source: Juniper Networks

source: Alcatel, Tyco

- State-of-the art up to 1 Tb/s throughput
- Lots of linecards – power constrained system
  - What matters is energy cost per bit

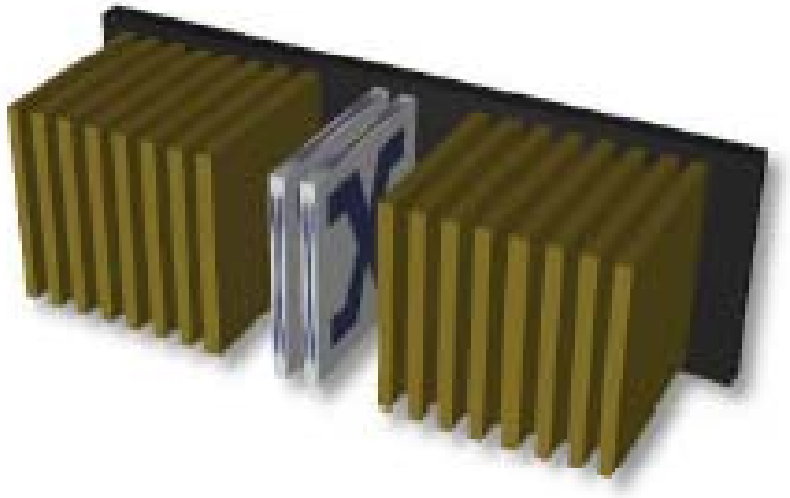
Inside the router



- Regardless of where the links are, there is a constant desire to signal faster and with less power

Scaling the throughput to 100 Tb/s

Electrical I/O Challenges

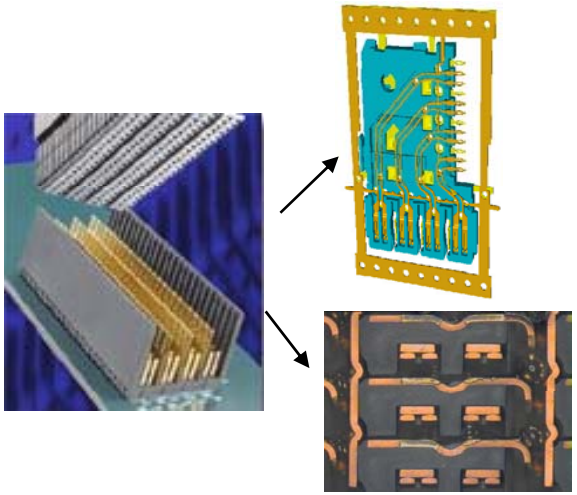


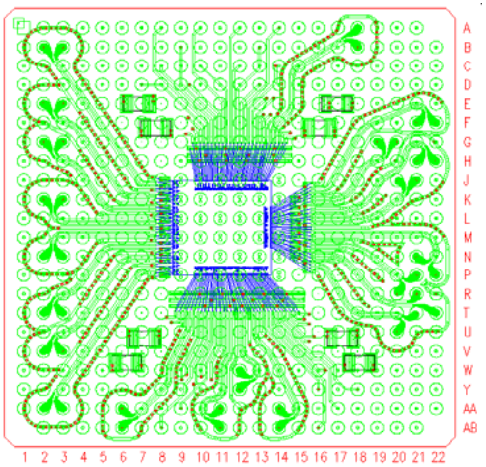
- 100 Tb/s I/O throughput
- With 10Gb/s per link
  - 10000 transceivers
    - 20000 high-speed I/O pairs
    - 10000 mm^2 in 0.13 µm techn
- Power 4kW
  - 40 mW/Gb/s – energy cost per bit

Scaling the throughput to 100 Tb/s

Density issues

- Connectors
  - 50 diff pairs/inch
  - 400” long connector
- Trace routing
  - 50mils pitch
  - 250” wide 4-signal layer line-card
  - Backplane less critical
- Package
  - Package/Chip ball pitch (1mm / 200um)
  - 4000 mm^2 / 160mm^2





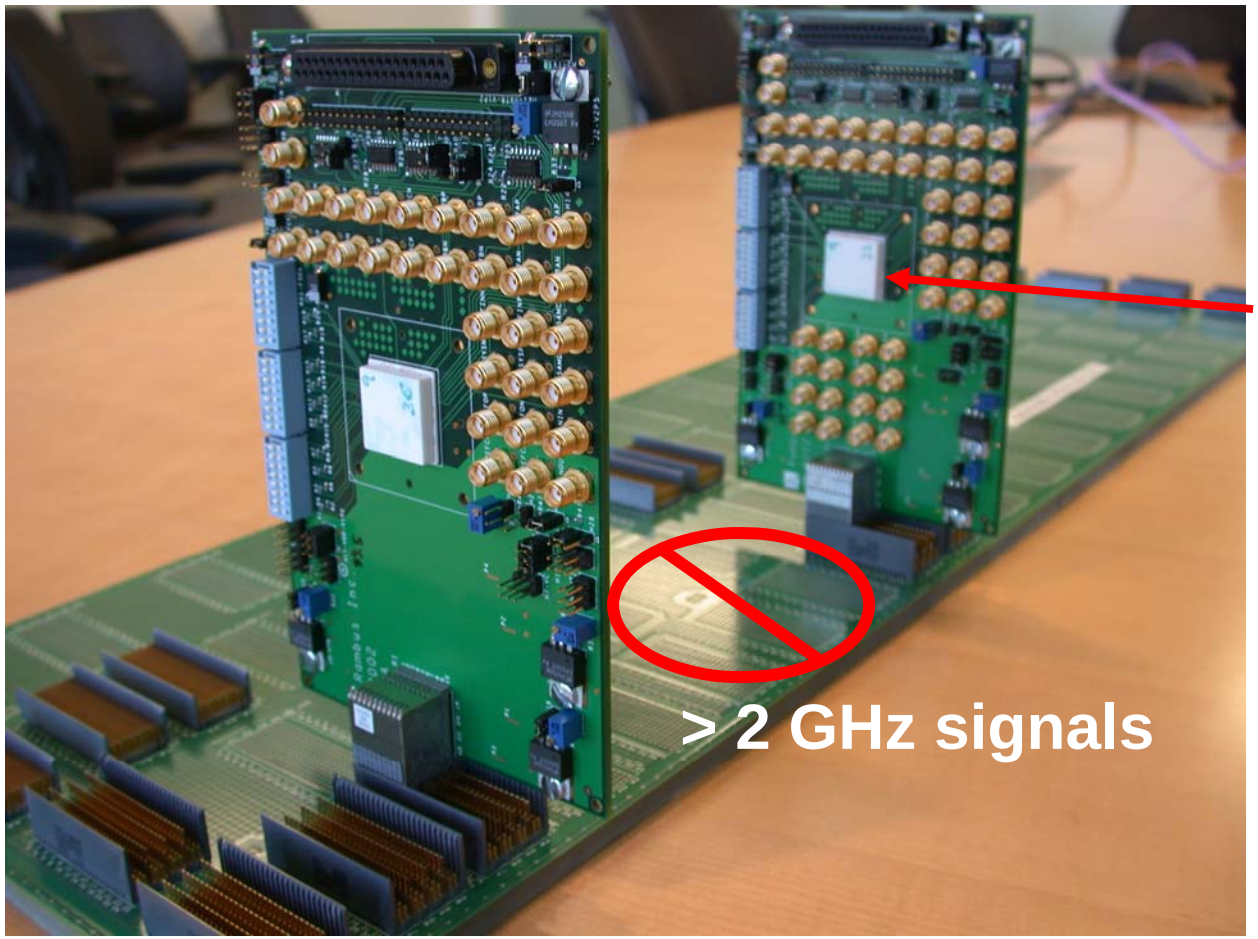
source: Teradyne, Rambus

Design challenge

- Goal
  - Fit 100 Tb/s on a 100 W crossbar chip
  - Reasonable system/rack size

- Need
  - Power
    - Reduce energy/bit to 1mW/Gb/s
  - Density
    - Increase data rate per link by 10-15x

What makes it challenging

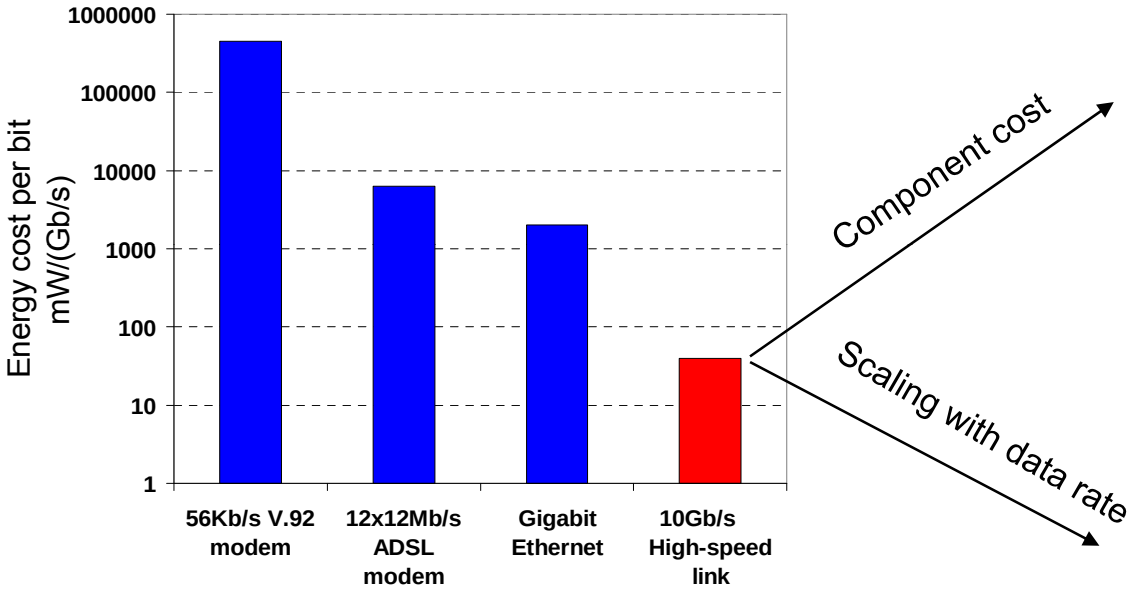


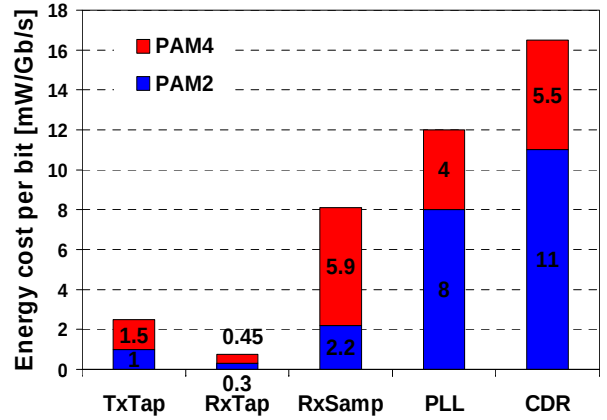
source: Rambus

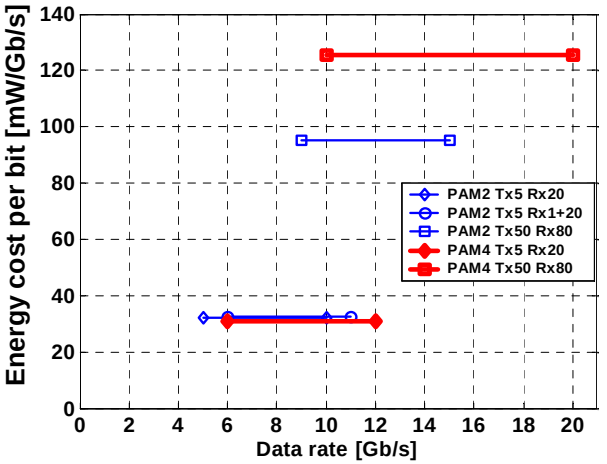
- Now, the bandwidth limit is in wires

High-speed link efficiency – energy cost per bit

How efficient are high-speed links?







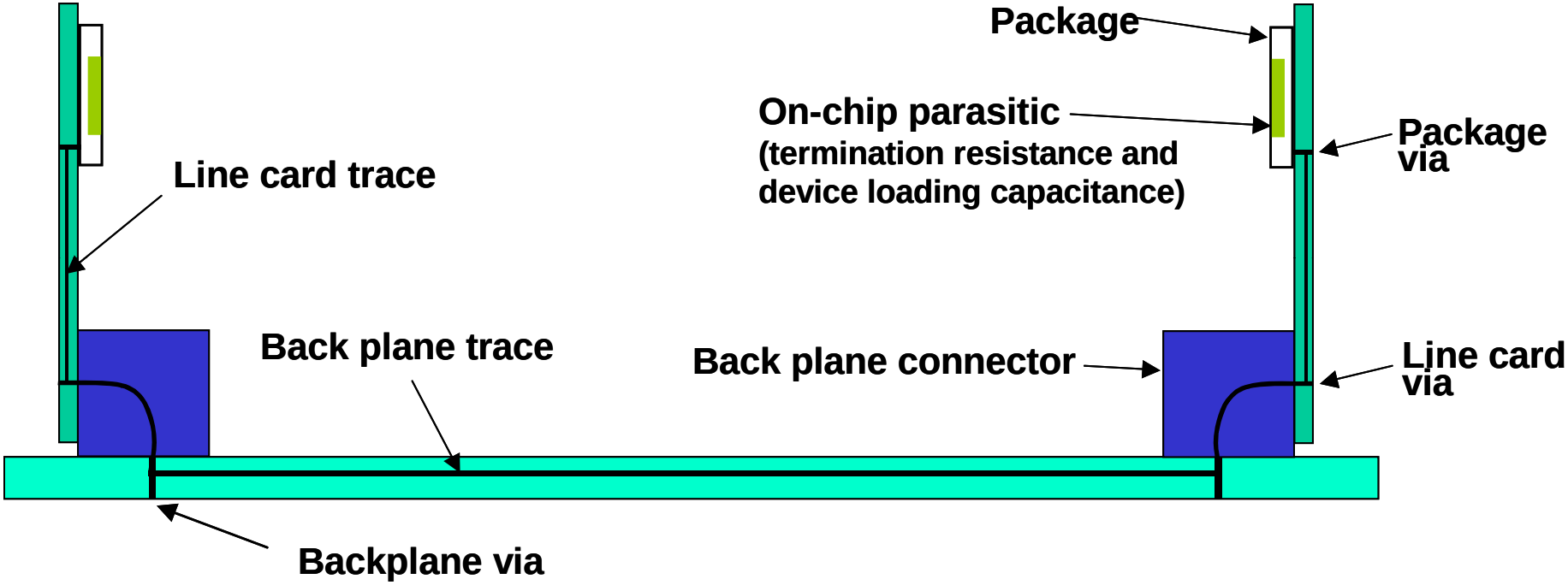
- 2-3 orders more energy-efficient
  - Than traditional wireline systems
- Starting to pay the price for band-limited channels

Outline

- Show the path to efficient 100 Tb/s systems
  - Look at all aspects of system design

- High-speed link environment
- Improving the channel
- What can chips do?

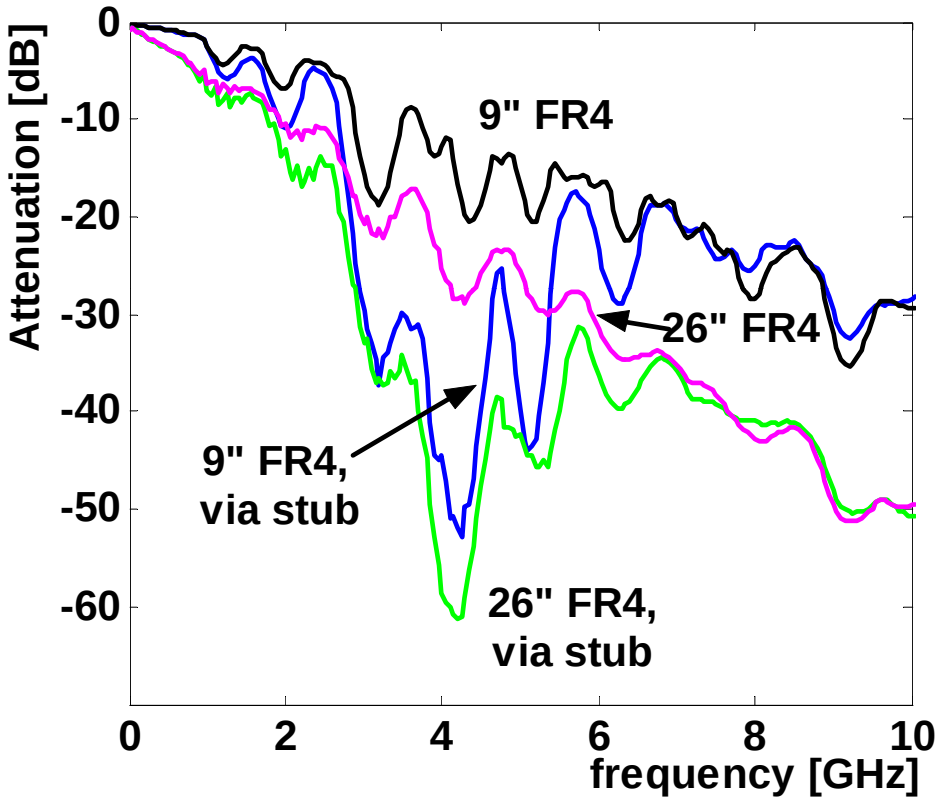
Backplane environment



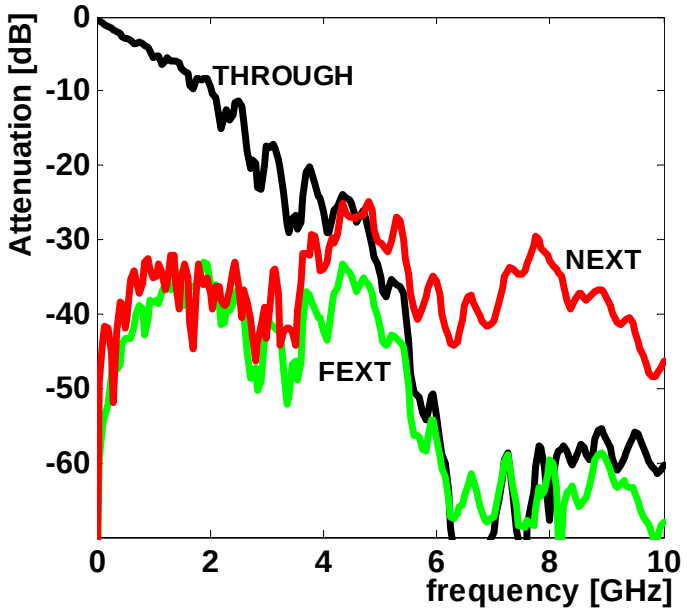
- Line attenuation
- Reflections from stubs (vias)

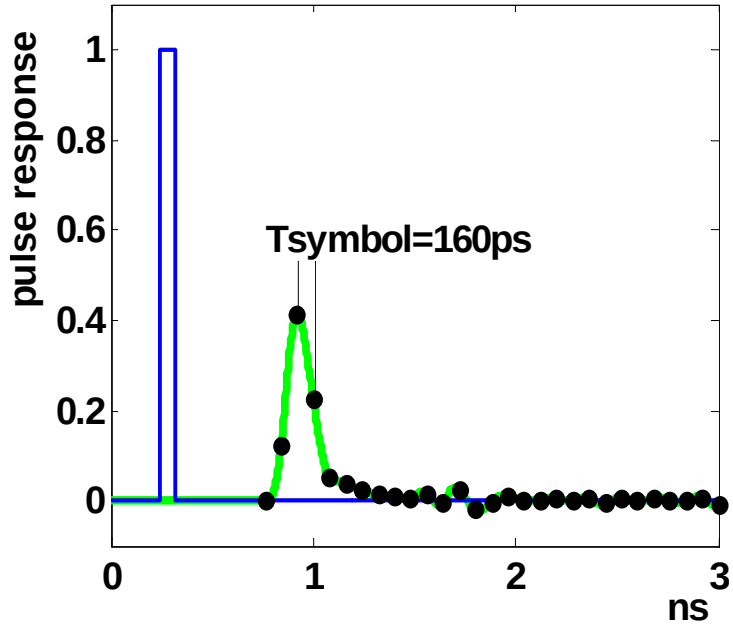
Backplane channel

- Loss is variable
  - Same backplane
  - Different lengths
  - Different stubs
    - Top vs. Bot
- Attenuation is large
  - >30dB @ 3GHz
  - But is that bad?
- Required signal amplitude set by noise



Interference

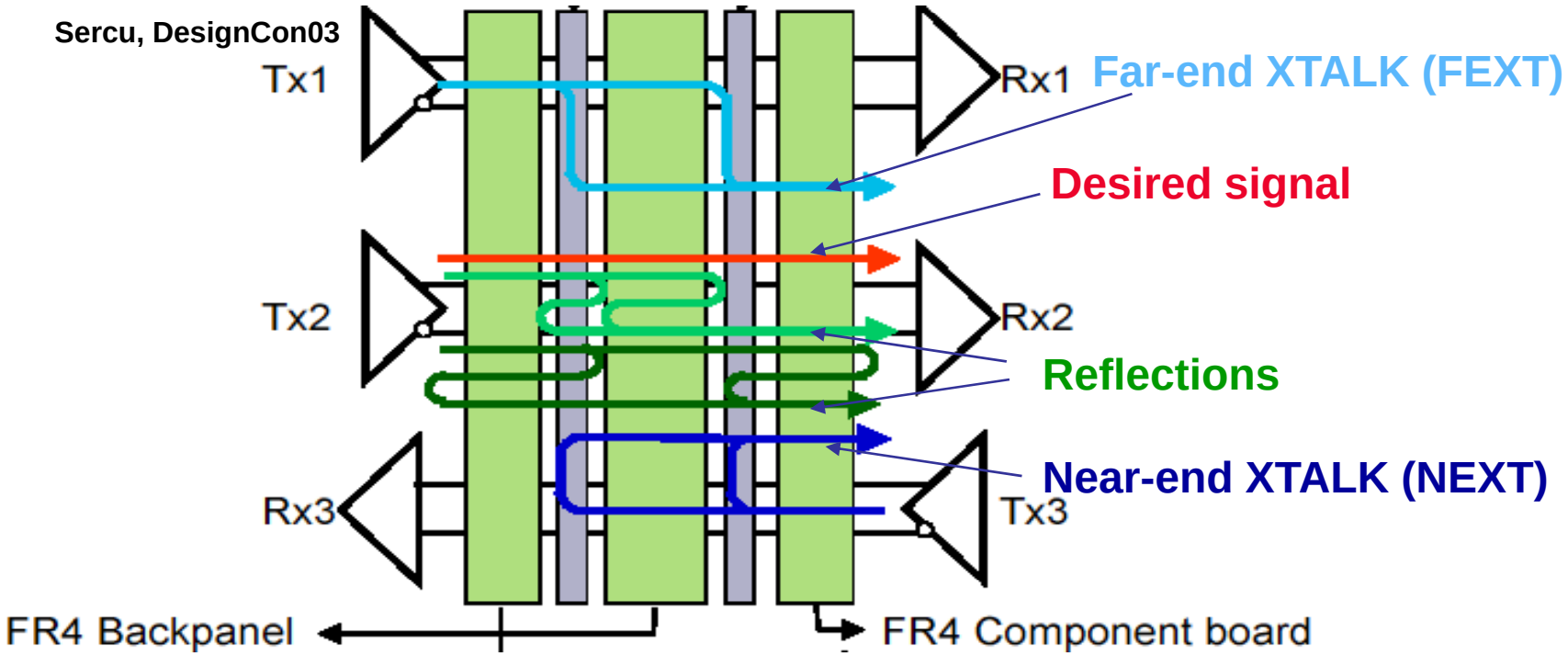




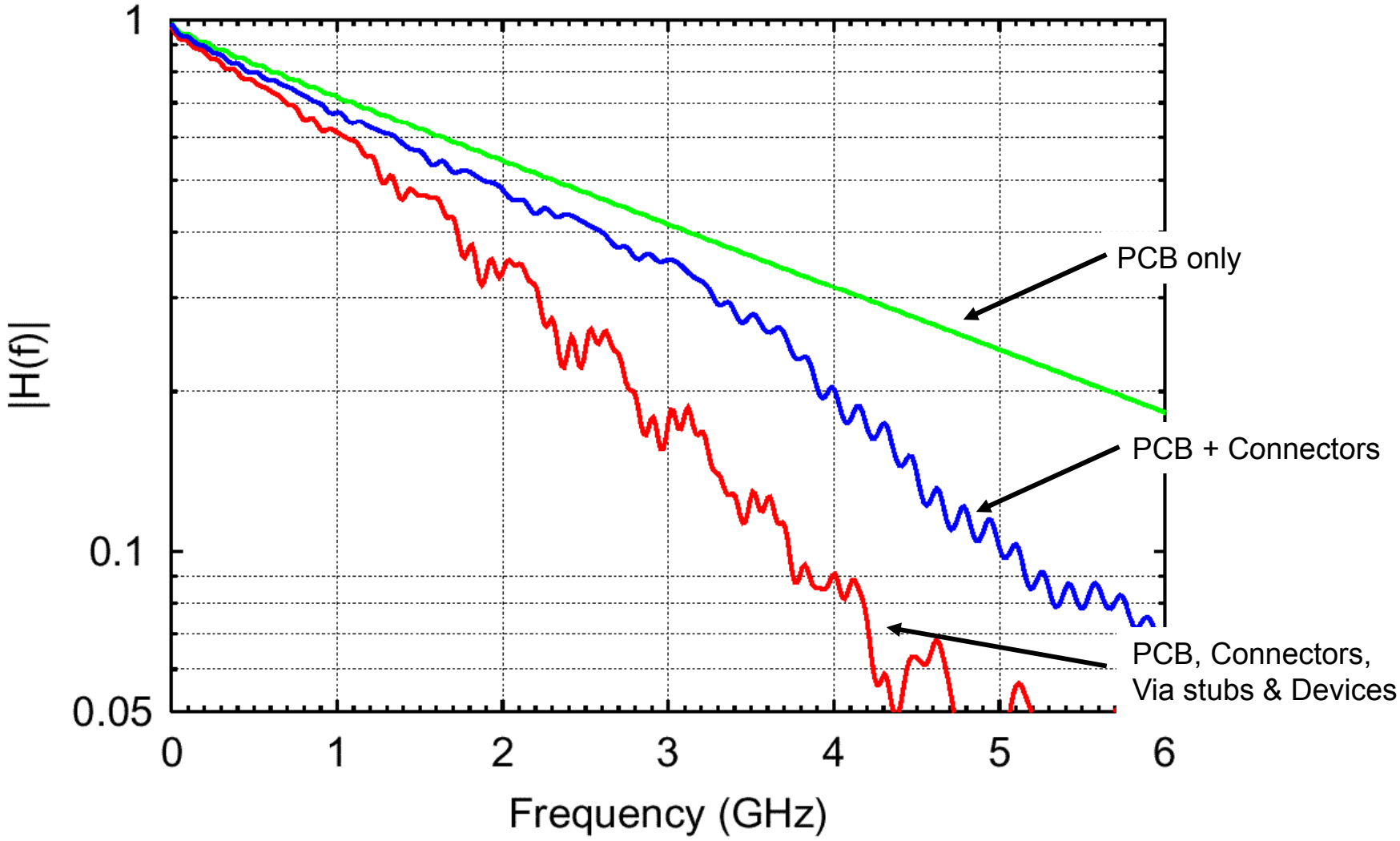
- Inter-symbol interference
  - Dispersion (skin-effect, dielectric loss) - short latency
  - Reflections (impedance mismatches – connectors, via stubs, device parasitics, package) – long latency
- Co-channel interference (Far-End & Near-End Crosstalk)

Reflections and Crosstalk

- Don't just receive the signal you want
  - Get versions of signals "close" to you
  - Vertical connections have worst coupling
    - "Close" in these vertical connection regions



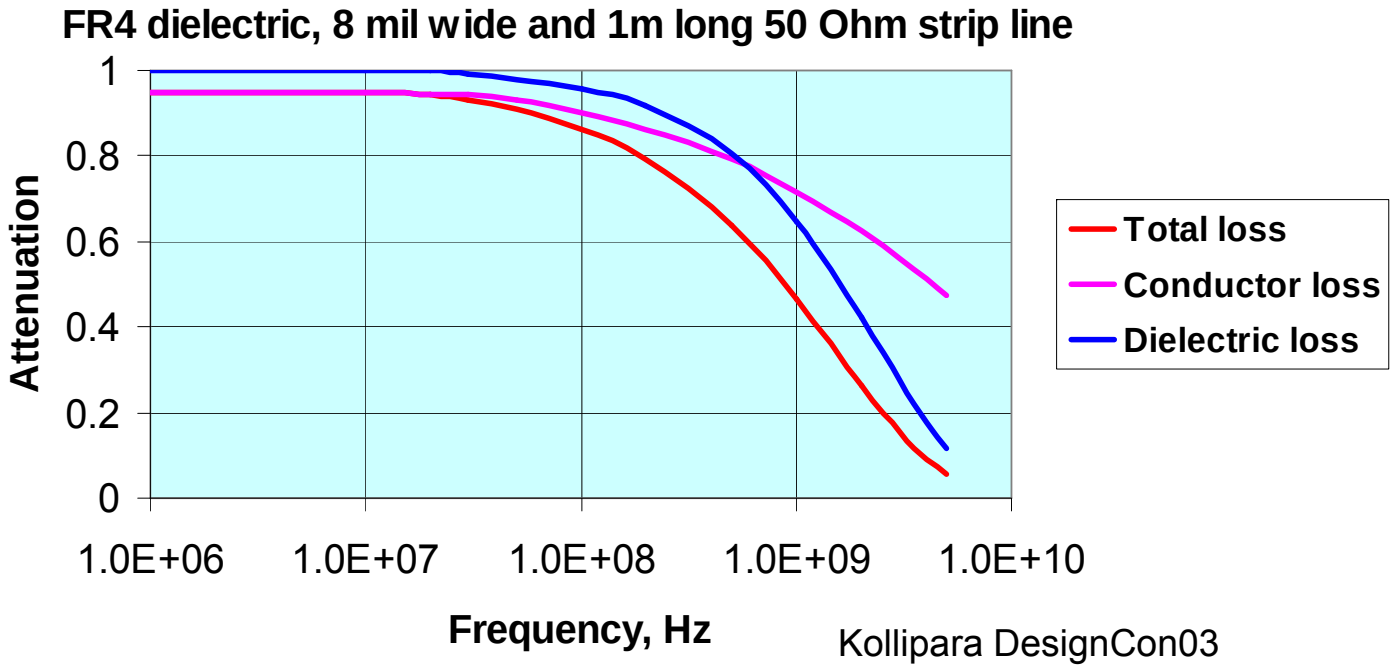
A complex system



Outline

- Show the path to efficient 100 Tb/s systems
  - Look at all aspects of system design
- High-speed link environment
- Improving the channel
- What can chips do?

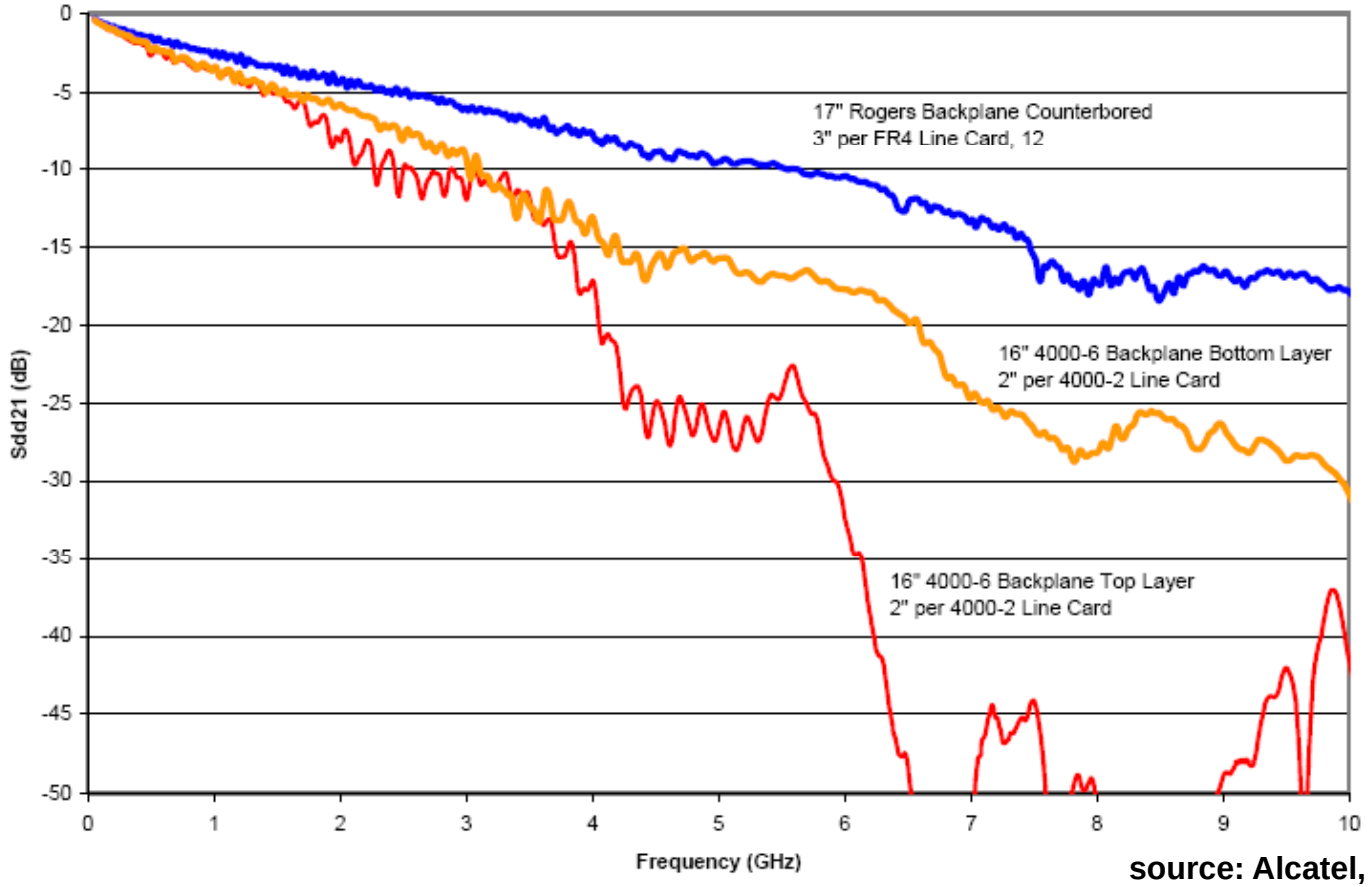
Dispersion: material loss



Kollipara DesignCon03

- PCB Loss : skin & dielectric loss
  - Skin Loss $\propto \sqrt{f}$
  - Dielectric loss $\propto$ f : a bigger issue at higher f

Better dielectric



Rogers

FR4

FR4+stubs

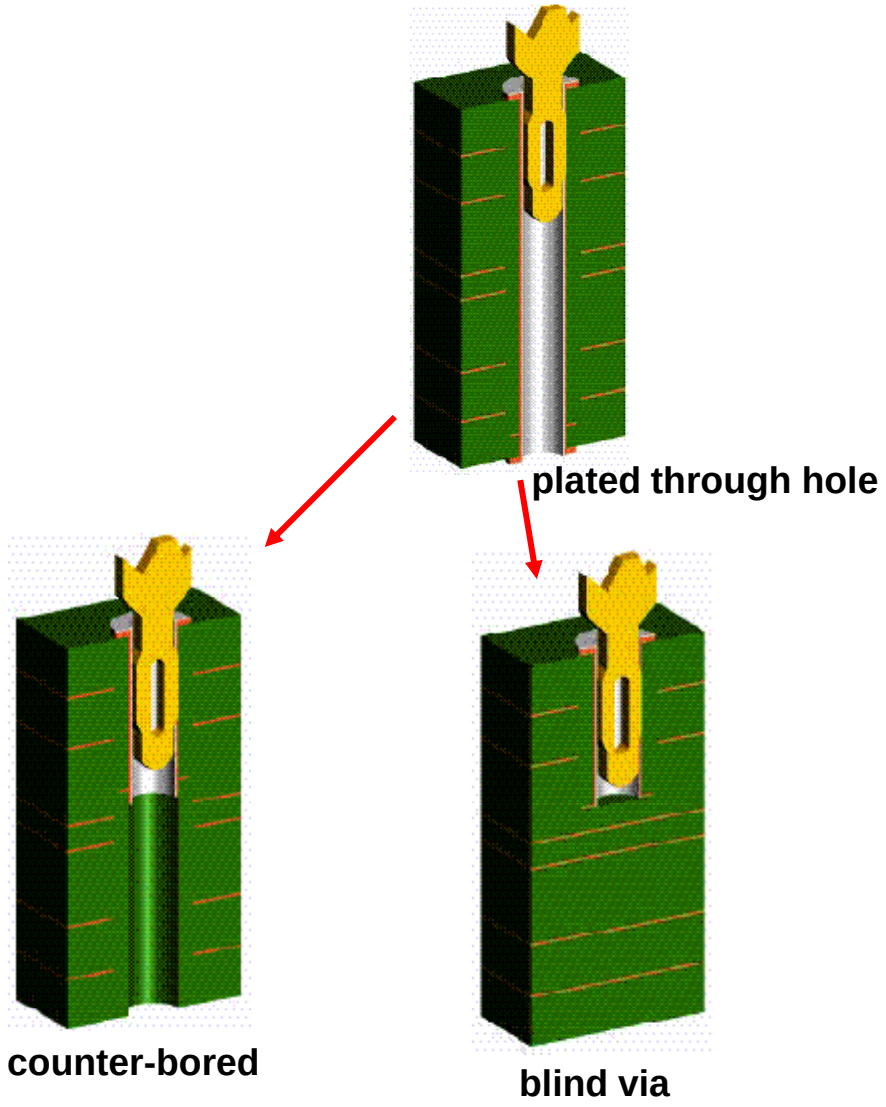
source: Alcatel, Tyco

- Rogers is expensive – but smallest loss

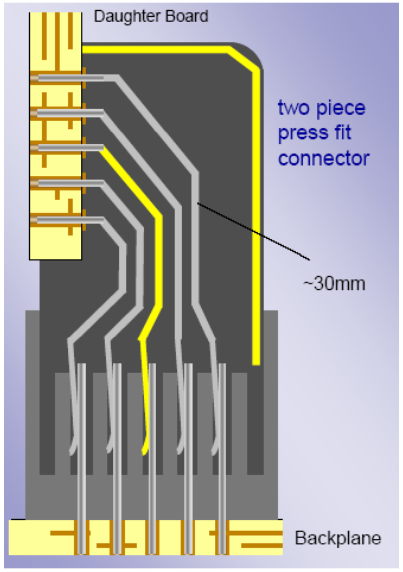
Minimizing reflections - the vias

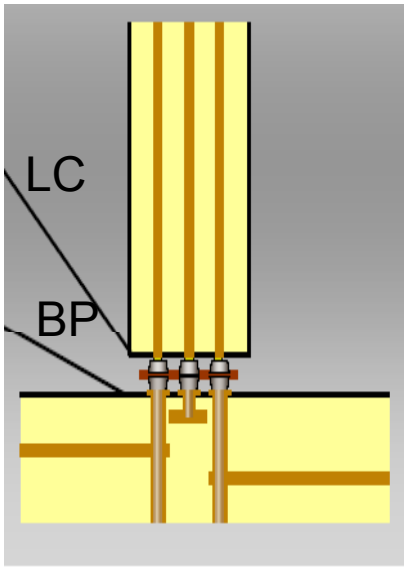
- Minimizing via stubs
  - Thinner PCBs are better…
    but sometimes impossible
  - Counter-boring
  - Blind vias
  - SMT technology

  - All are costly
    1.1x - 2x

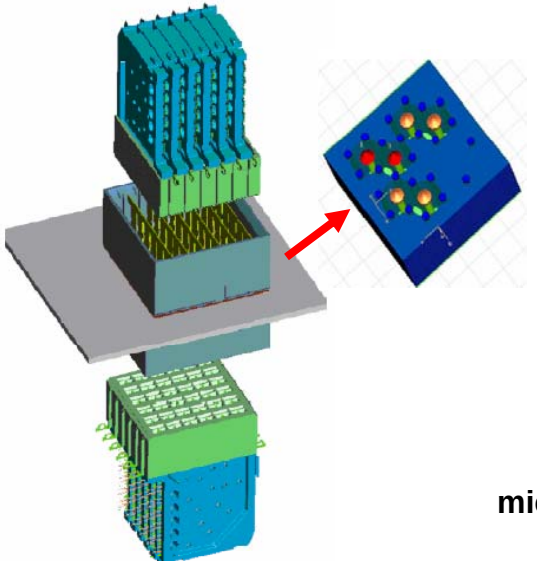


Connector technologies

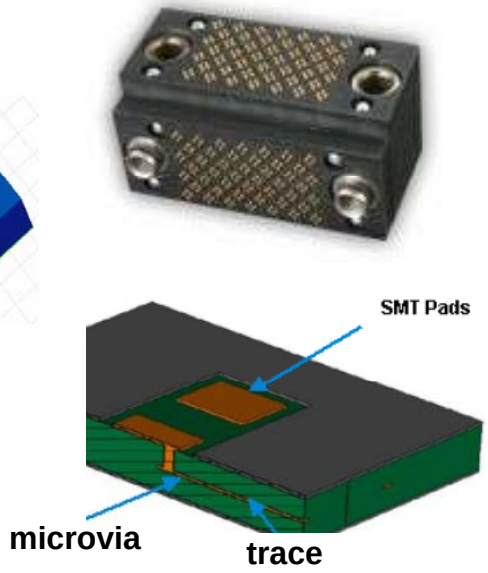




Standard - Press-Fit Side-Interface (Tyco)



Orthogonal - Teradyne
(Differential Plated Through Hole)

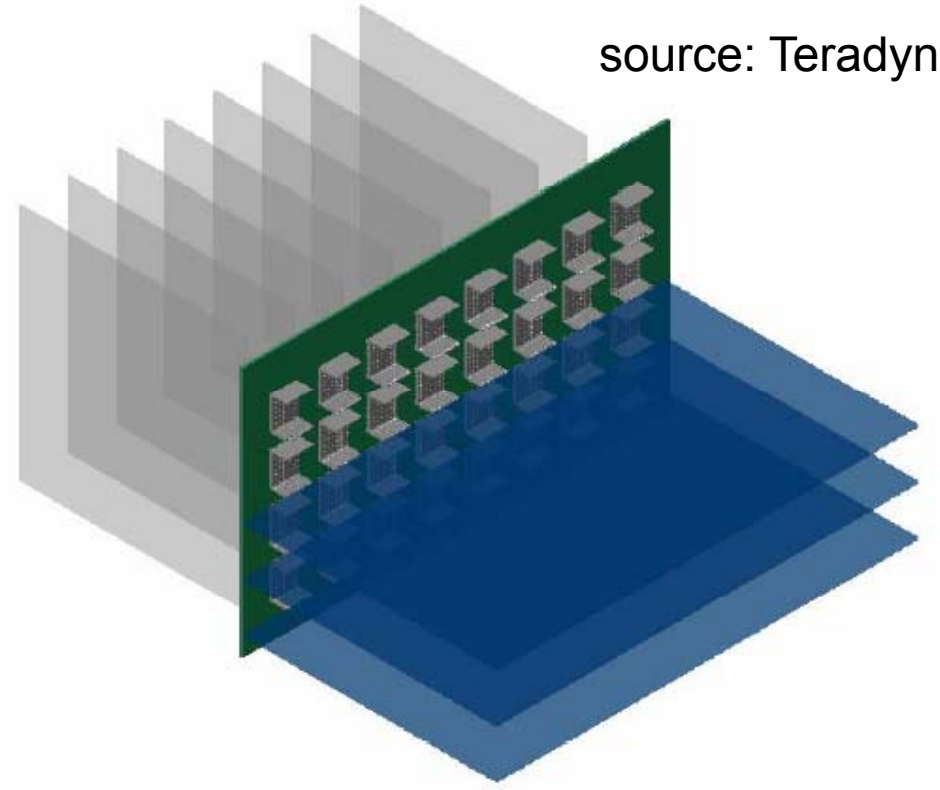


Surface-mount + microvia

- Stubs big problem in standard press-fit connectors
- Side-Interface eliminates DC stubs and diff-pair length mismatch
- Orthogonal interconnect DPTH eliminates the backplane
- Surface-mount

Eliminating the backplane - orthogonal interconnect

source: Teradyne



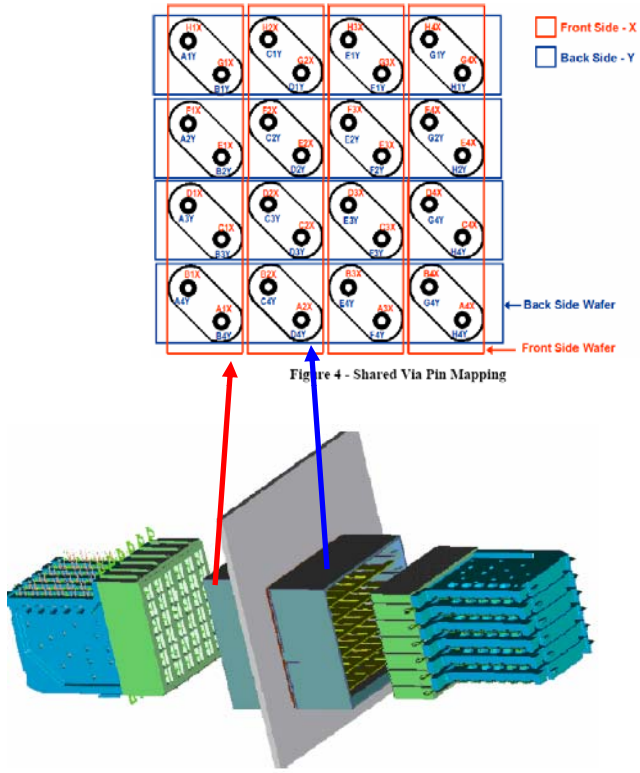


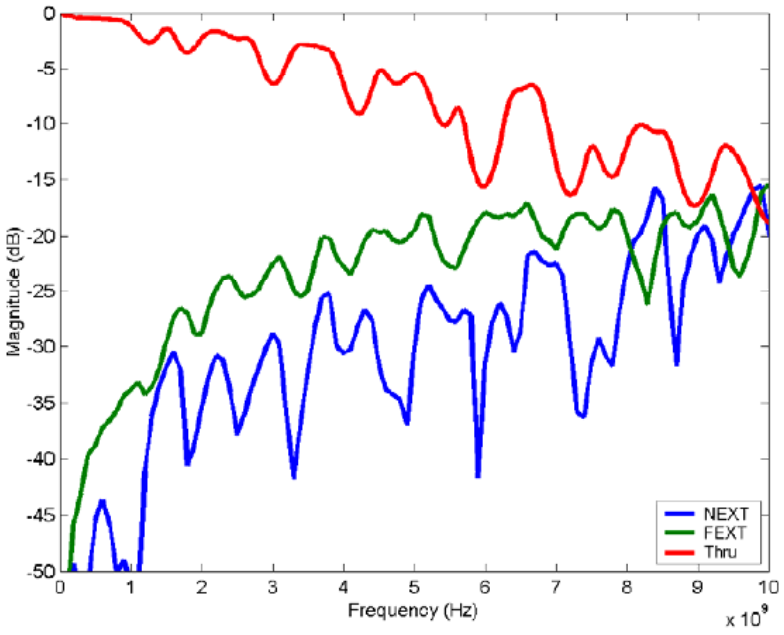
Figure 4 - Shared Via Pin Mapping

- No backplane trace
- No backplane via-stub
- Coax-like shielding and diff-pair matching in DPTH mid-plane

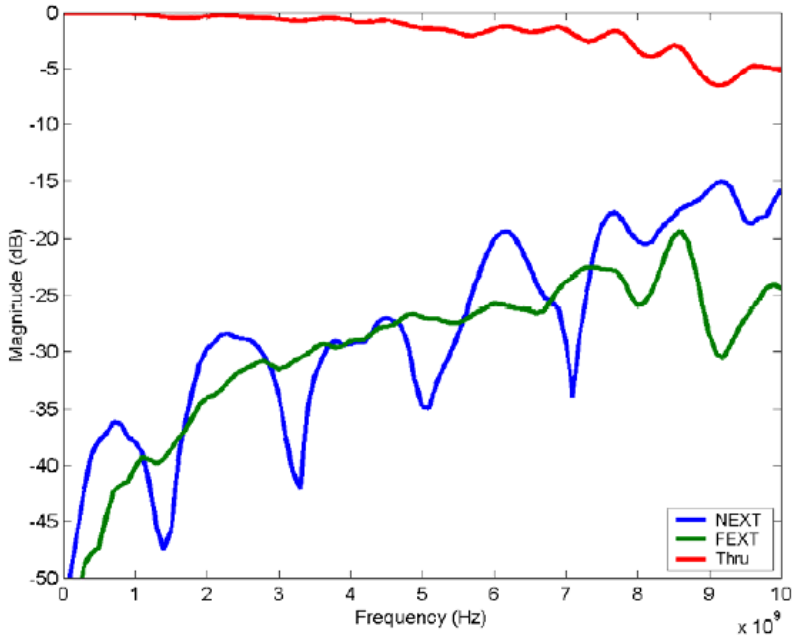
M. Cartier et al "Optimized Signal Path for Orthogonal System Architectures," DesignCon 2005.

DPTH connector performance

No shared vias (non-DPTH)



Shared vias (DPTH)



- Insertion Loss of DPTH very small
- Reflections minimized
- NEXT and FEXT minimized

Outline

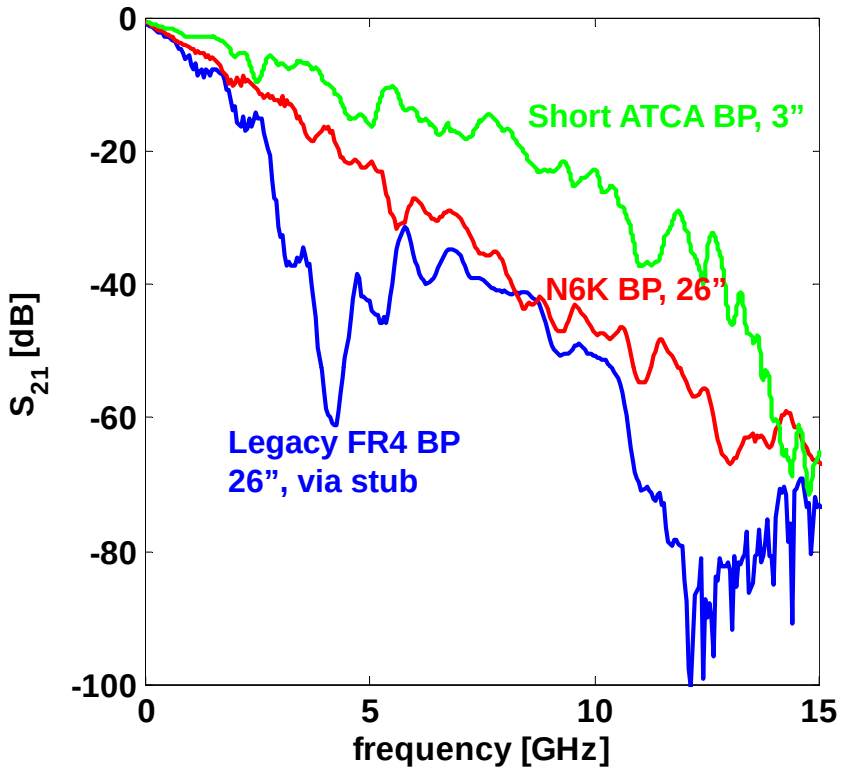
- Show the path to efficient 100Tb/s systems
  - Look at all aspects of system design
- High-speed link environment
- Improving the channel
- What can chips do?

New link design

Dealing with bandwidth limited channels

- This is an old research area
  - Textbooks on digital communications
  - Think modems, DSL
- But can't directly apply their solutions
  - Standard approach requires high-speed A/Ds and digital signal processing
  - 20Gs/s A/Ds are expensive
- (Un)fortunately need to rethink issues

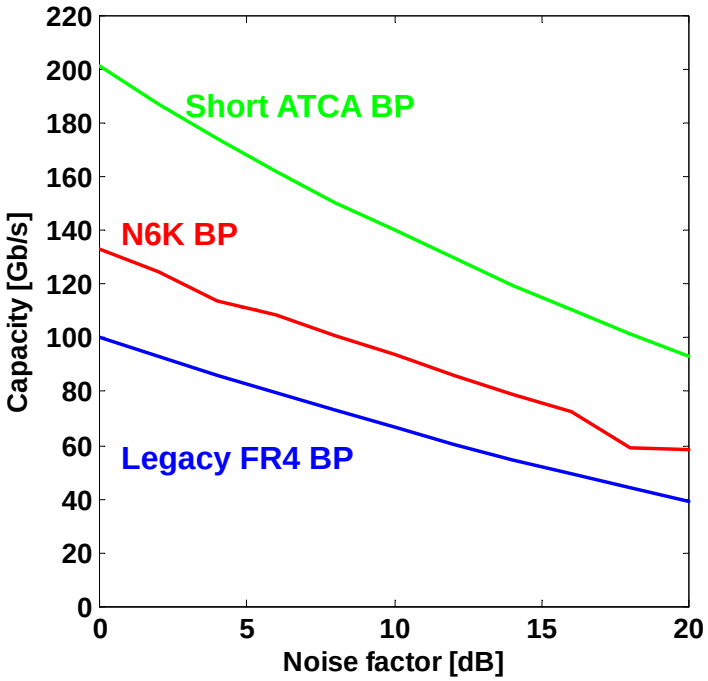
Baseline Channels



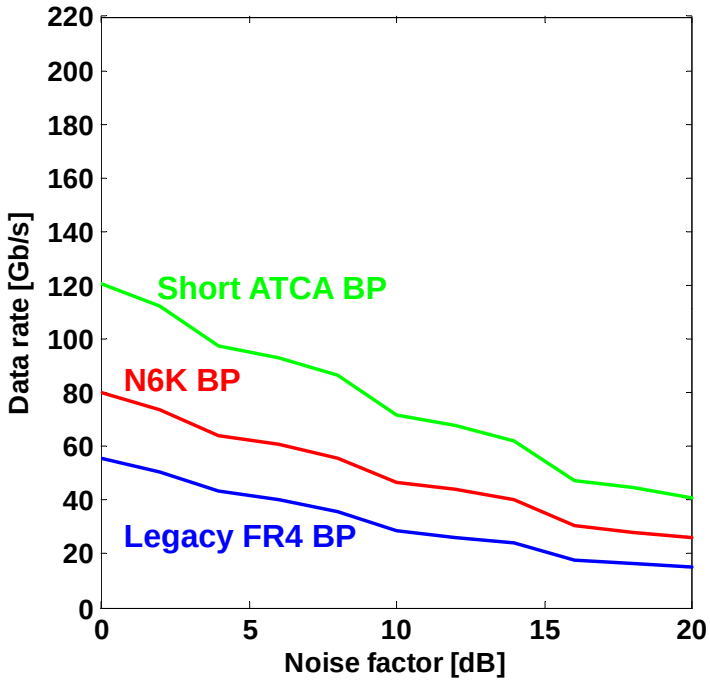
- Legacy (FR4) - lots of reflections
- Microwave engineered (N6K)
- Emerging standards (IEEE 802.3ap, ATCA)

Capacity and MT data rates – the impact of noise

Capacity – thermal and phase noise



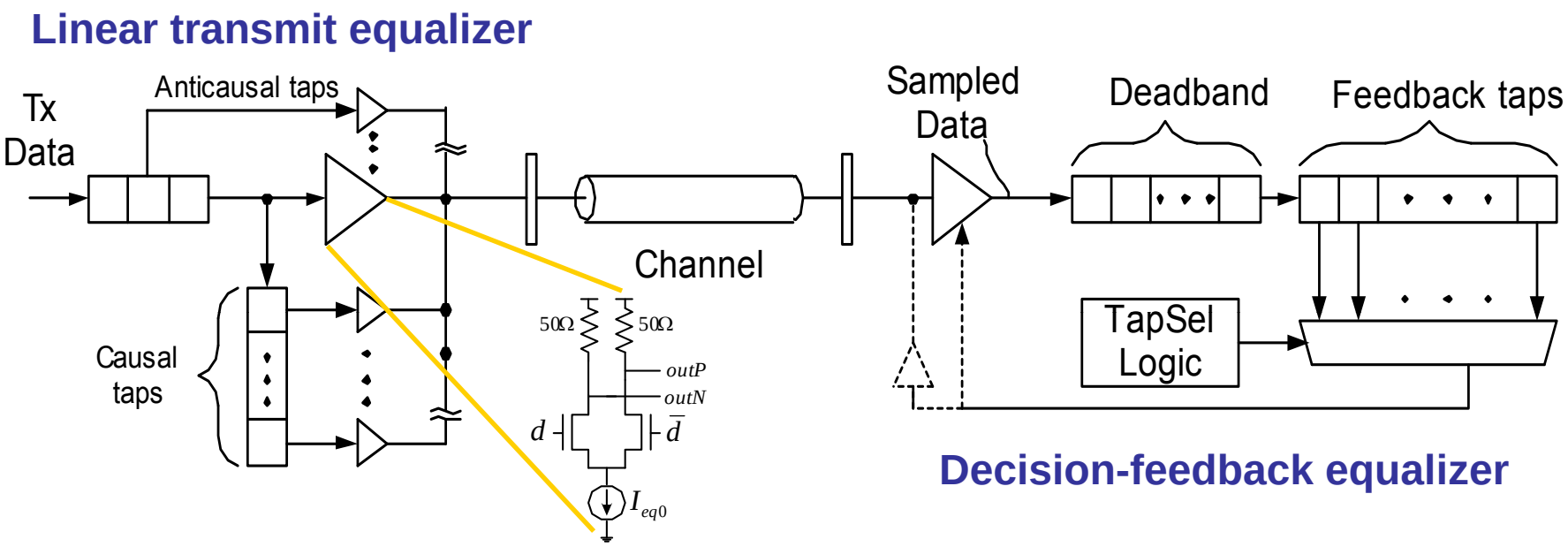
Uncoded MT – thermal and phase noise



- Capacity
  - Much higher than data rates in today's links
  - Noise
    - Thermal - 50Ohm termination
    - Phase noise – best LC PLL (0.14%UI rms)

- Uncoded MT
  - Half the capacity
    - BER target of 10^{-15}
    - Peak-power constraint
  - Coding can help

Removing ISI – baseband link

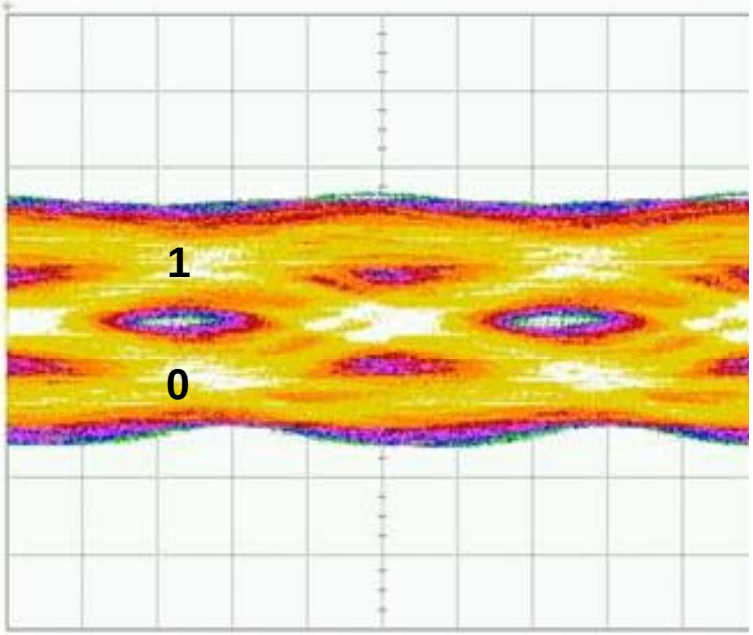


- Transmit and Receive Equalization
  - Changes signal to correct for ISI
  - Often easier to work at transmitter
    - DACs easier than ADCs

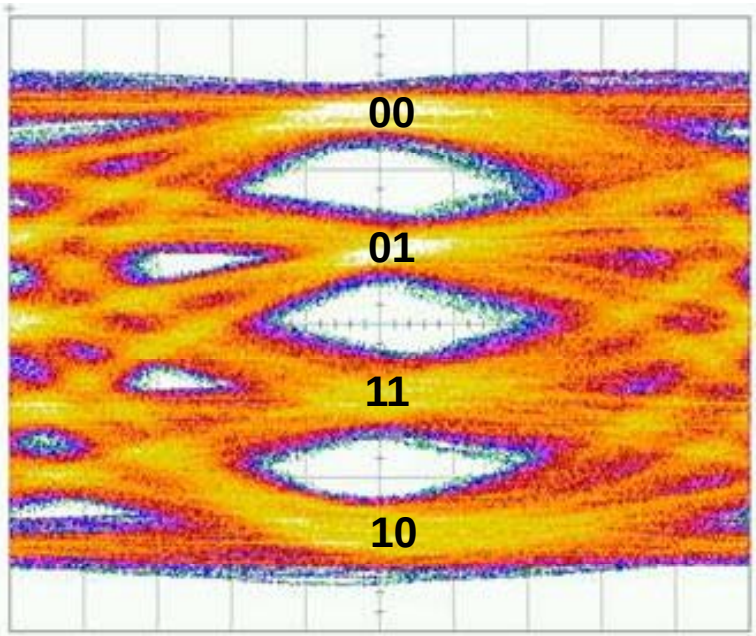
J. Zerbe et al, "Design, Equalization and Clock Recovery for a 2.5-10Gb/s 2-PAM/4-PAM Backplane Transceiver Cell," *IEEE Journal Solid-State Circuits*, Dec. 2003.

Pulse amplitude modulation

- Binary (NRZ)
  - 1 bit / symbol
  - Symbol rate = bit rate

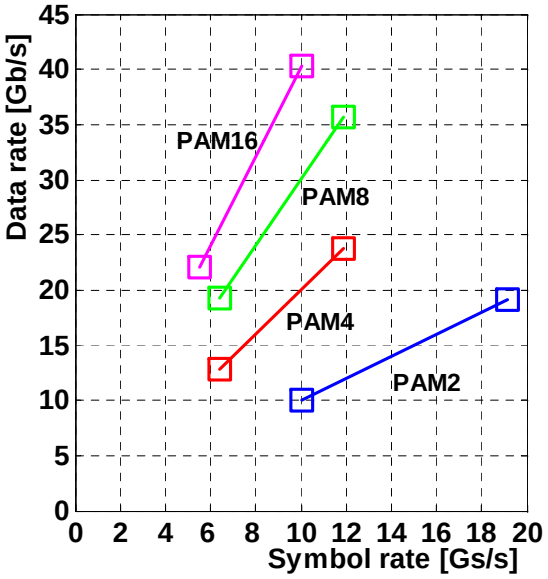


- PAM4
  - 2 bits / symbol
  - Symbol rate = bit rate/2

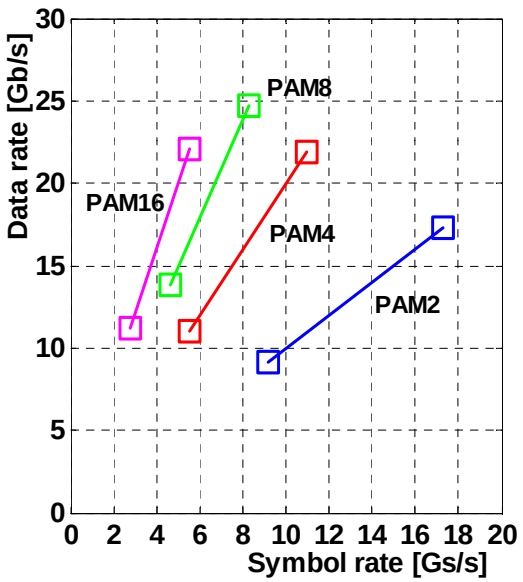


Multi-level: offset and jitter are crucial

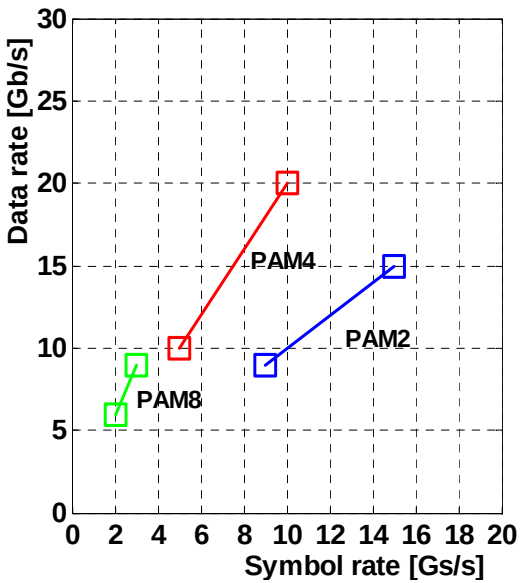
thermal noise



thermal noise + offset



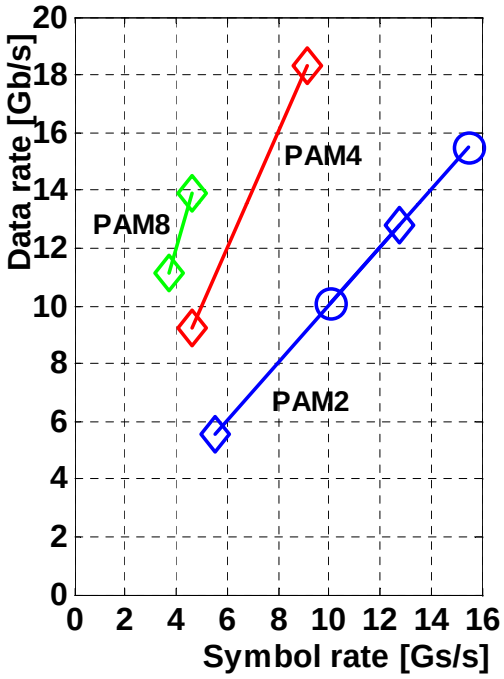
thermal noise + offset+ jitter



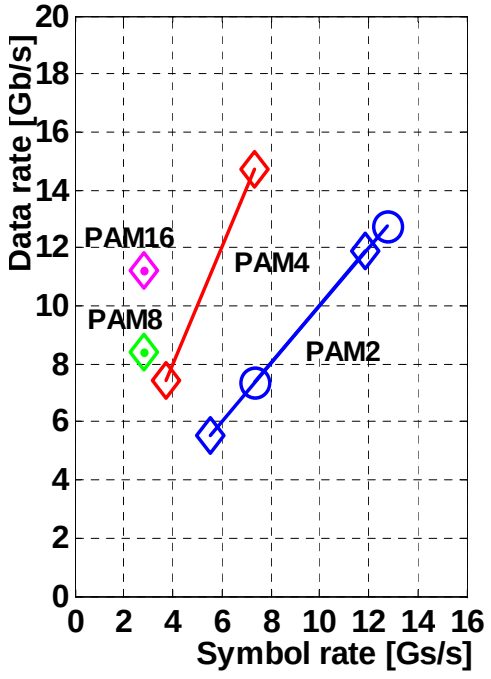
- To make better use of available bandwidth, need better circuits
- PAM2/PAM4 robust candidate for next generation links

Full ISI compensation too costly

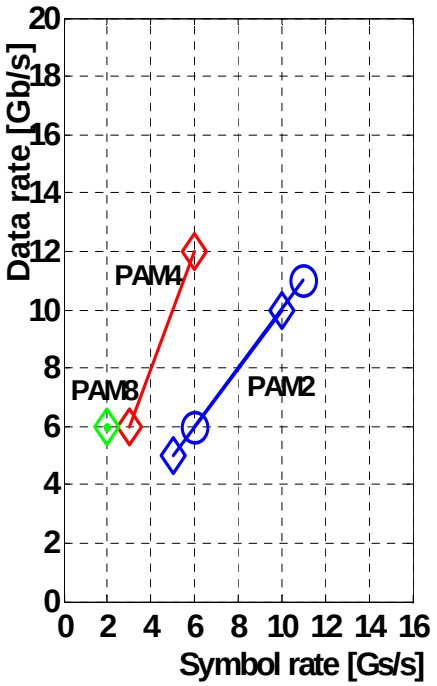
thermal noise



thermal noise + offset

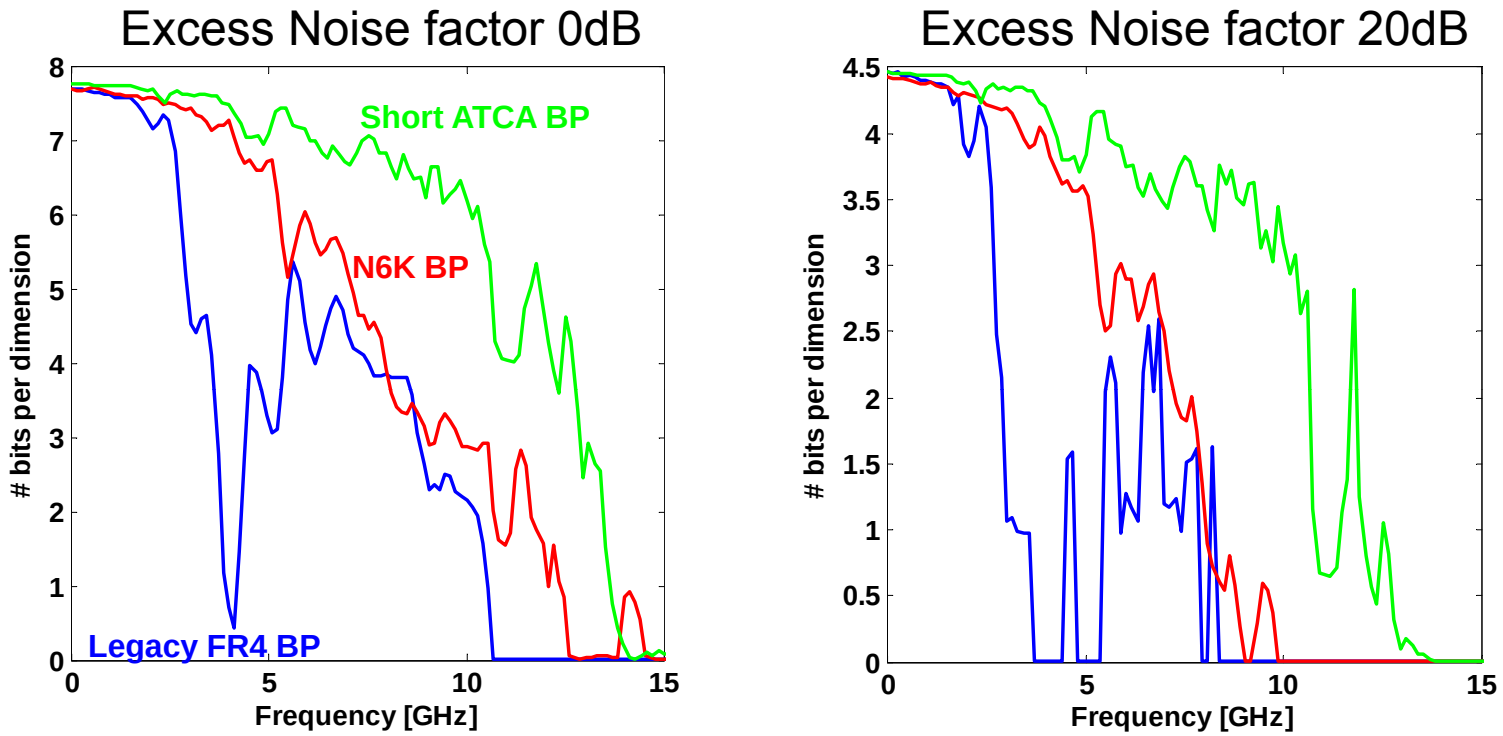


thermal noise + offset+ jitter



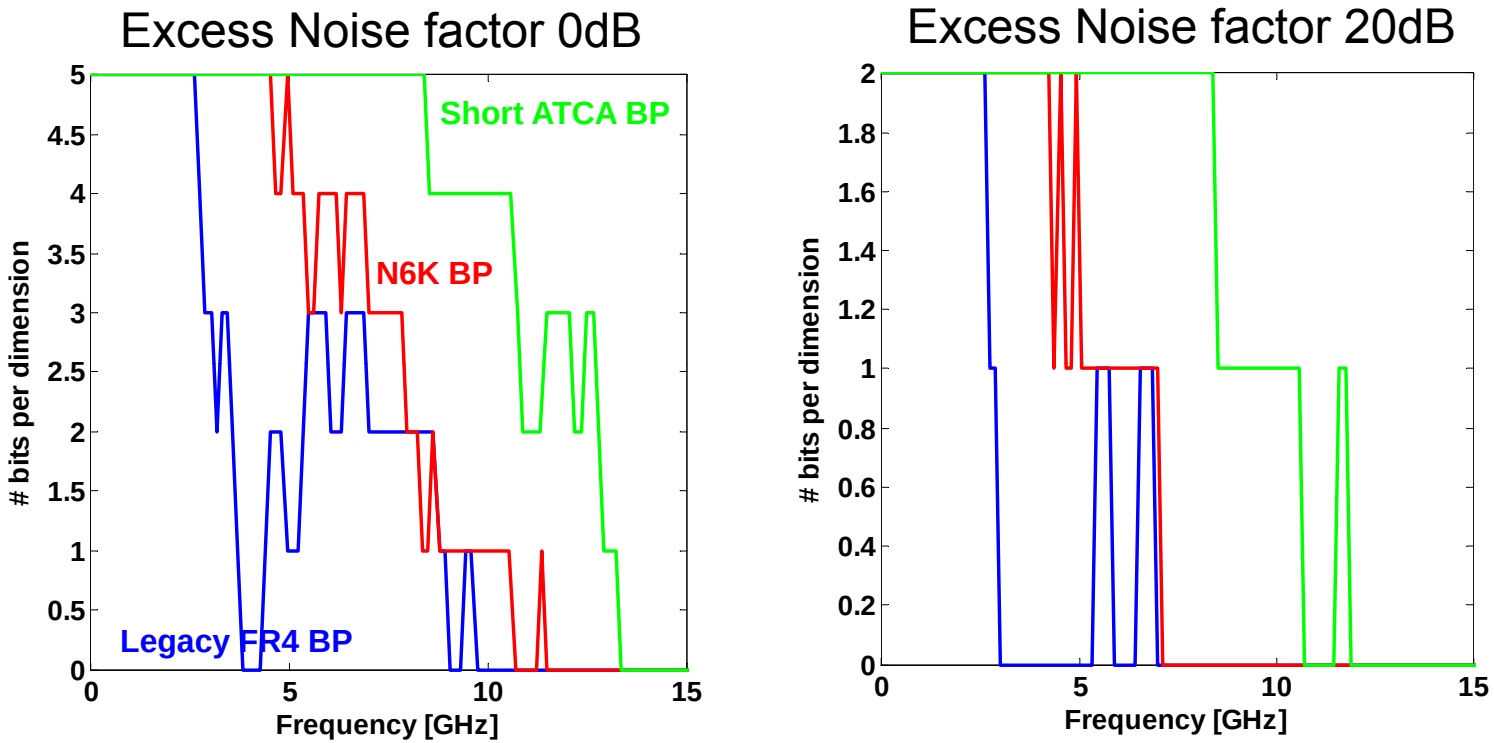
- Today's links cannot afford to compensate all ISI
  - Too much power
  - Limits today's maximum achievable data rates

Capacity – Bit Loading



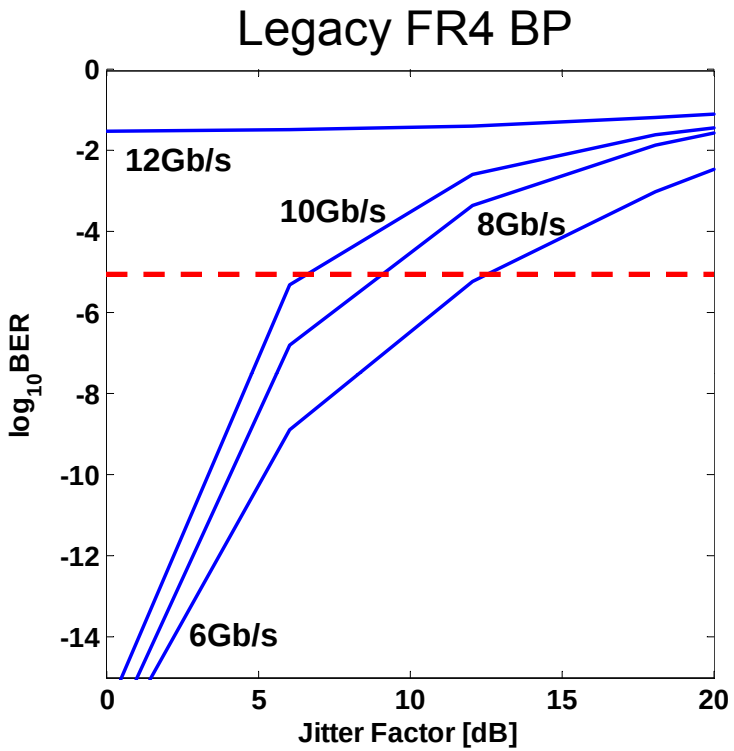
- Bandwidth is limited by attenuation and noise
  - Can't just keep increasing the signaling frequency
  - Need to focus on available bandwidth (at most 10-20GHz)
- Need circuits that can create/sense 4-8 bits/dim

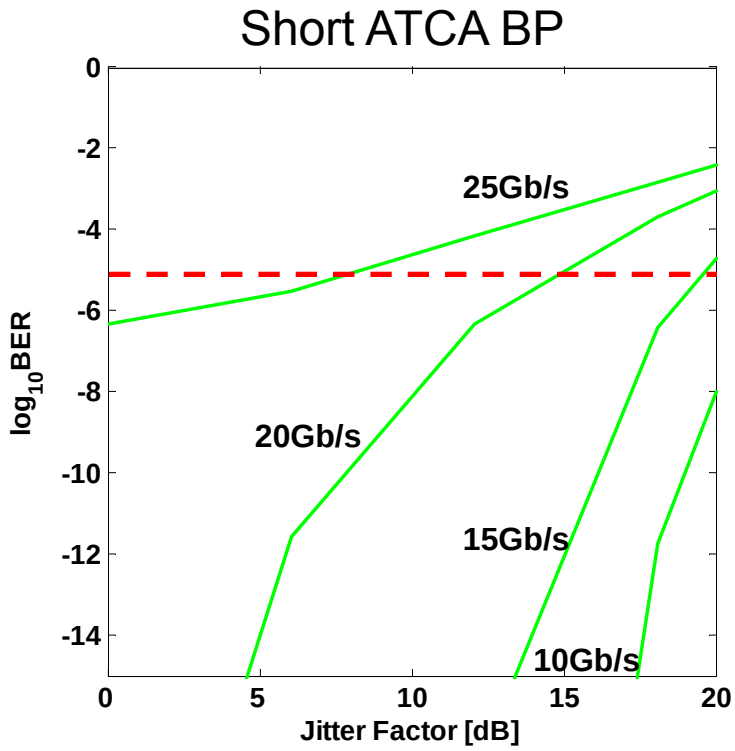
Uncoded Multi-tone – Bit Loading



- Integer constellations and target BER=10^{-15}
  - Bandwidth not affected much (still 10-20GHz)
- In high-noise case - less advantage over baseband
- With coding can improve by up to 2x – closer to capacity

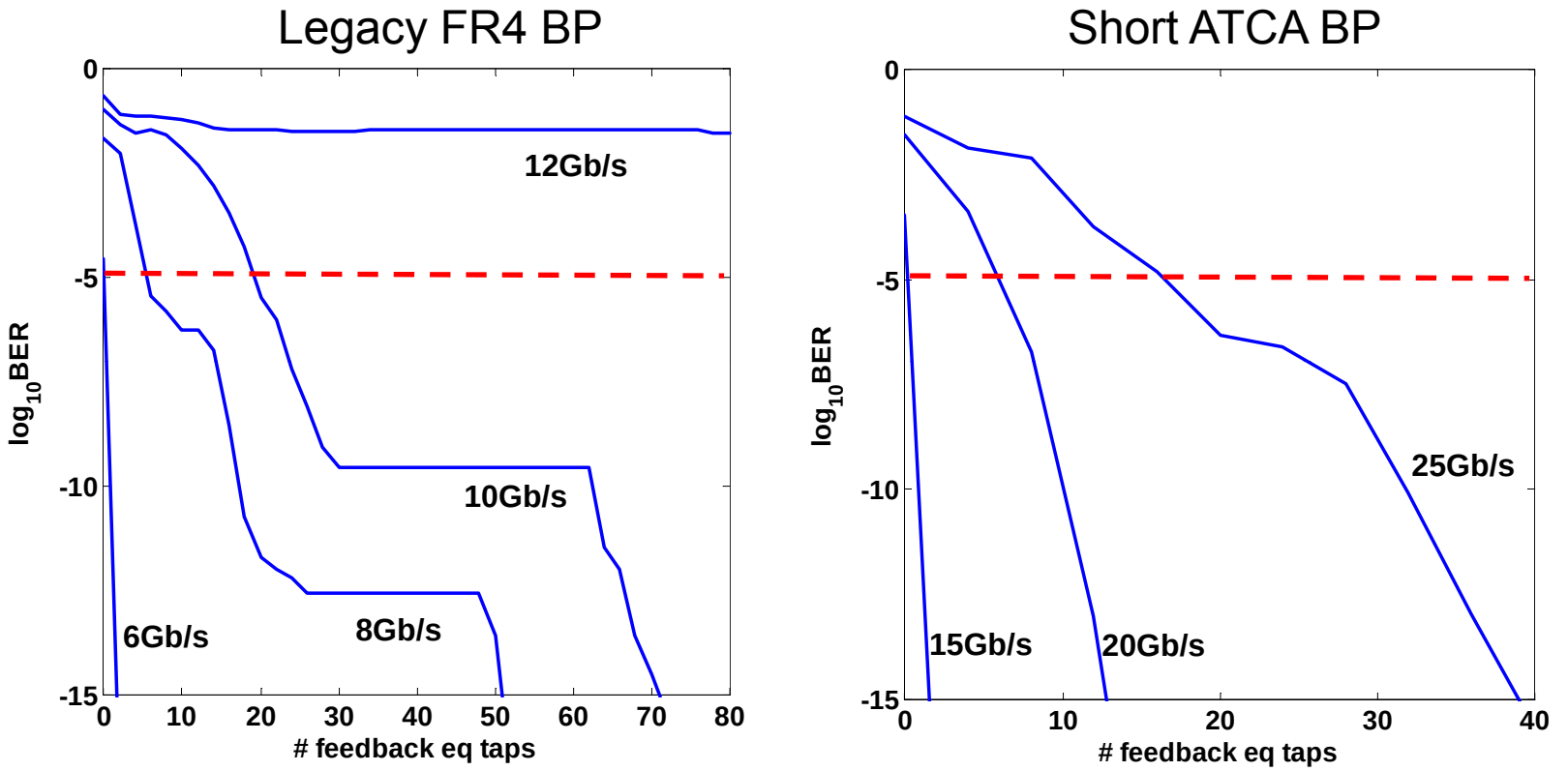
Impact of jitter on baseband





- With proper coding
  - Increase data rate
  - Relax PLL jitter spec – save power
- Original jitter – rms = 1.4%UI (ring oscillator based PLL)

BER vs. hardware complexity



- Partially eliminate ISI (leave most of the reflections)
- Let simple code take care of the rest
  - Can recover from raw BER of 10^{-5}
  - And save up to 50 feedback taps - up to 15mW/Gb/s in 0.13µm

But, need to be careful

- Always now what you're optimizing
  - Powerful coders/encoders often costly
    - Example - fastest RS (255,239) implementation
    - 10 – 40 Gb/s throughput
    - Energy cost - 12mW/Gb/s
    - 50x area of the high-speed link (extensive parallelism)

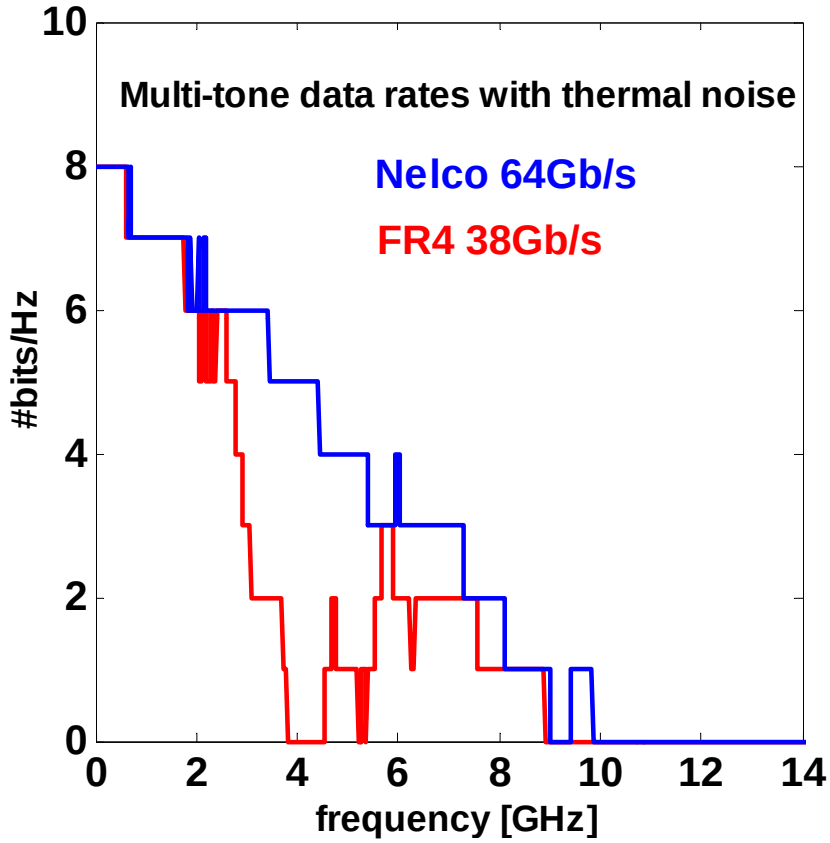
- Need to include the energy cost per bit in the code design spec

L. Song, M-L Yu, M.S. Shaffer, "10- and 40-Gb/s Forward Error Correction Devices for Optical Communications," ***IEEE Journal of Solid-State Circuits*****, vol. 37, no. 11, Nov. 2002.**

Opportunity for coding

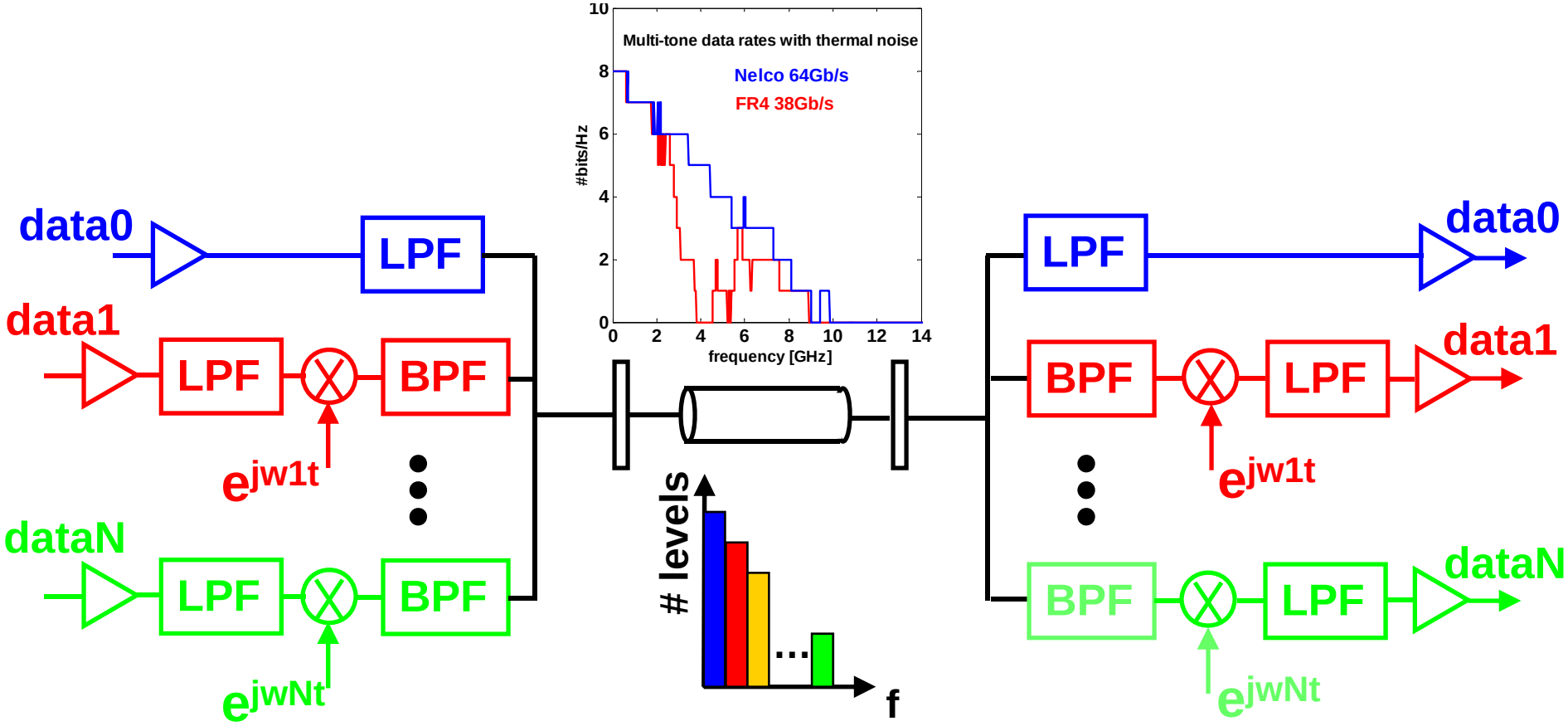
- Break the coding/equalization/modulation hierarchy
- Goal to minimize overall energy cost per bit
- Proper coding can be more energy-efficient in achieving the low BER than modulation/equalization
  - Especially with lots of crosstalk and numerous small reflections
- Need new paradigms in code development to specification
  - Non-Gaussian (system) noise
  - Circuit non-idealities
  - Crosstalk and residual channel memory (ISI)
  - Energy cost constraint on code performance

Bridging the gap: Multi-tone link



A. Amirkhany, V. Stojanovic, M.A. Horowitz, "Multi-tone Signaling for High-speed Backplane Electrical Links," *IEEE Global Telecommunications Conference*, November 2004.

Bridging the gap: Multi-tone link



- Challenge – balancing the inter-symbol and inter-channel interference
  - Microwave filter techniques
  - Custom signal processing

Conclusions

- Interfaces are challenging system designs
  - Good space to explore system level optimization
- Better backplanes are around the corner
  - 2-3x improvement in data rate possible
- State-of-the-art baseband links (chips)
  - Far from utilizing the capacity of the channels
    - 10-20x difference in data rates
  - Looking into multi-tone and coding to bridge the gap
  - Useful channel bandwidth 10-20 GHz
    - Focus on lower-speed precision circuits for higher order constellations
- Coding
  - If careful, can lower the energy cost per bit for the whole system
  - Problem formulation different in so many ways

Acknowledgments

- MARCO Interconnect Focus Center

- Jared Zerbe and Ravi Kollipara - Rambus
- John D'Ambrosia – Tyco
- IEEE 802.3ap, ATCA forum
- Alcatel, Teradyne, Juniper Networks