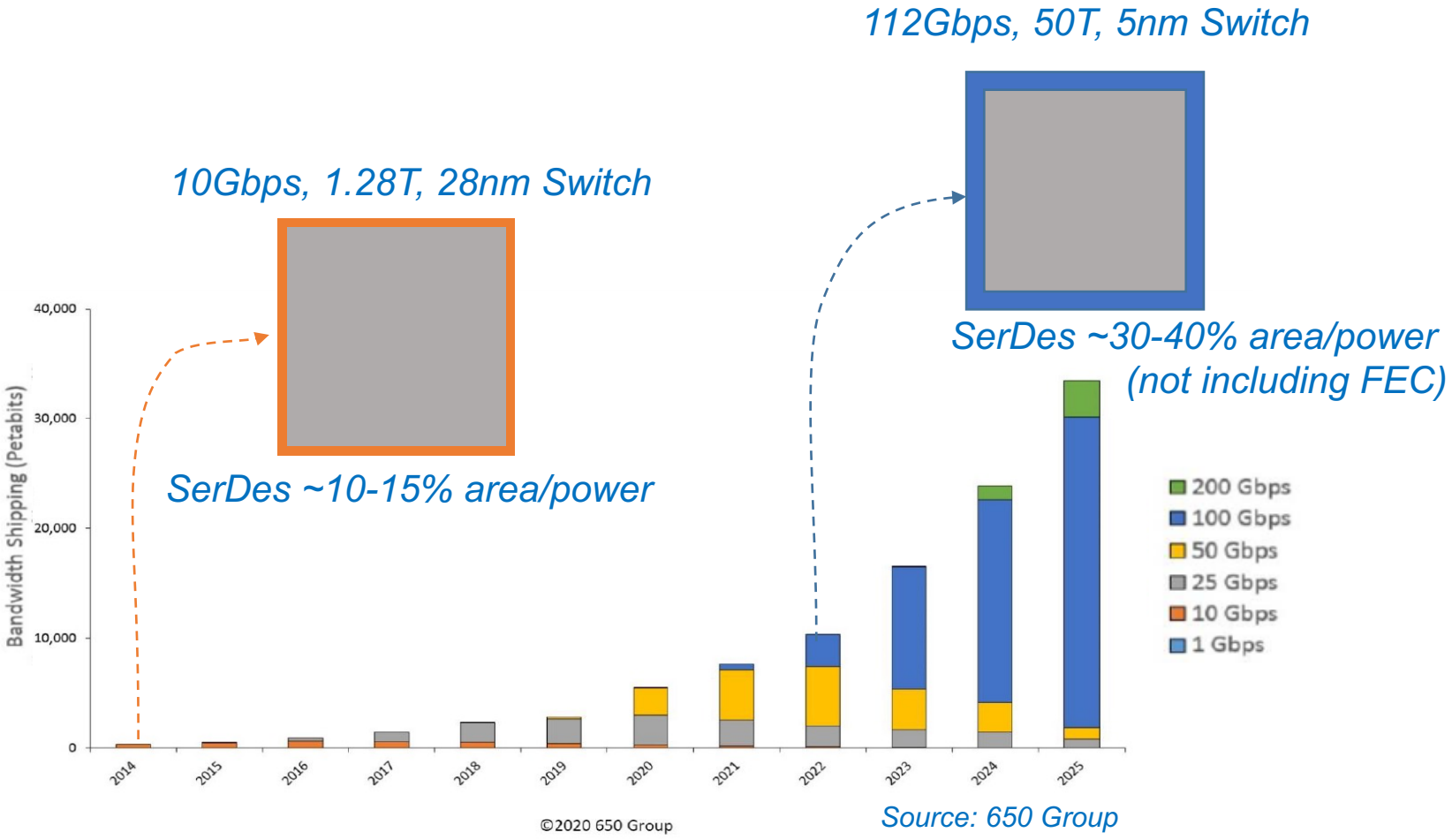
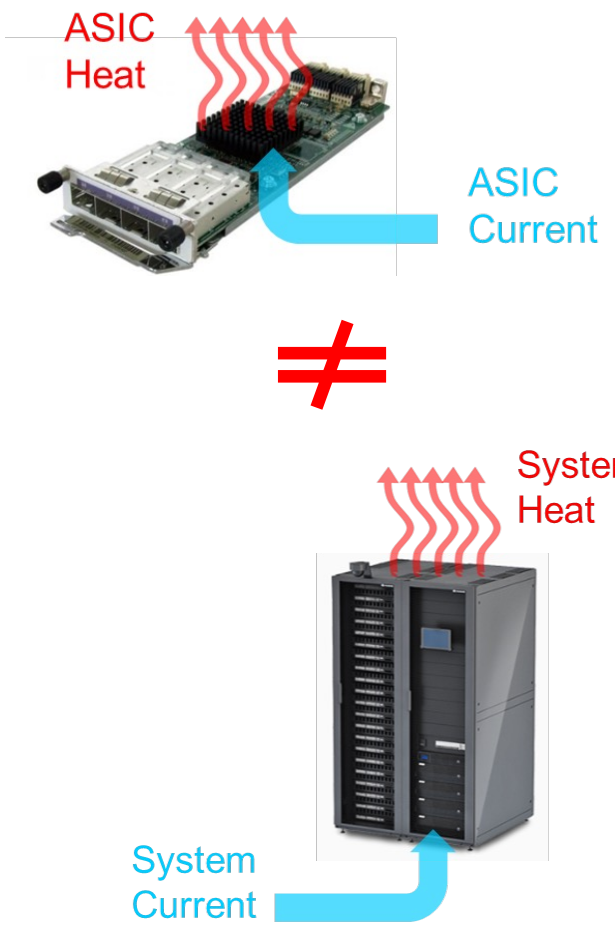
Energy Efficiency in Serial Links

Davide Tonietto
Huawei Fellow
Huawei Technologies Canada
davide.tonietto@huawei.com

Exponential Growth

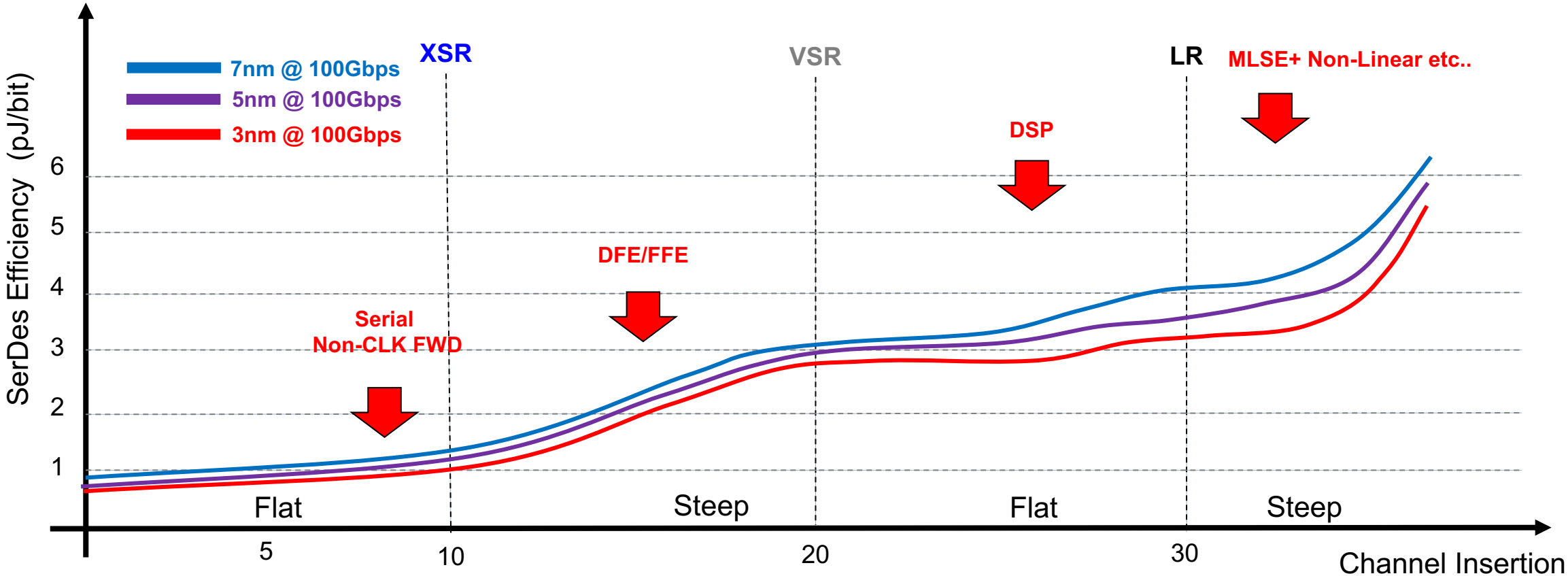


Total port bandwidth shipped per year



Energy Efficiency vs. Process

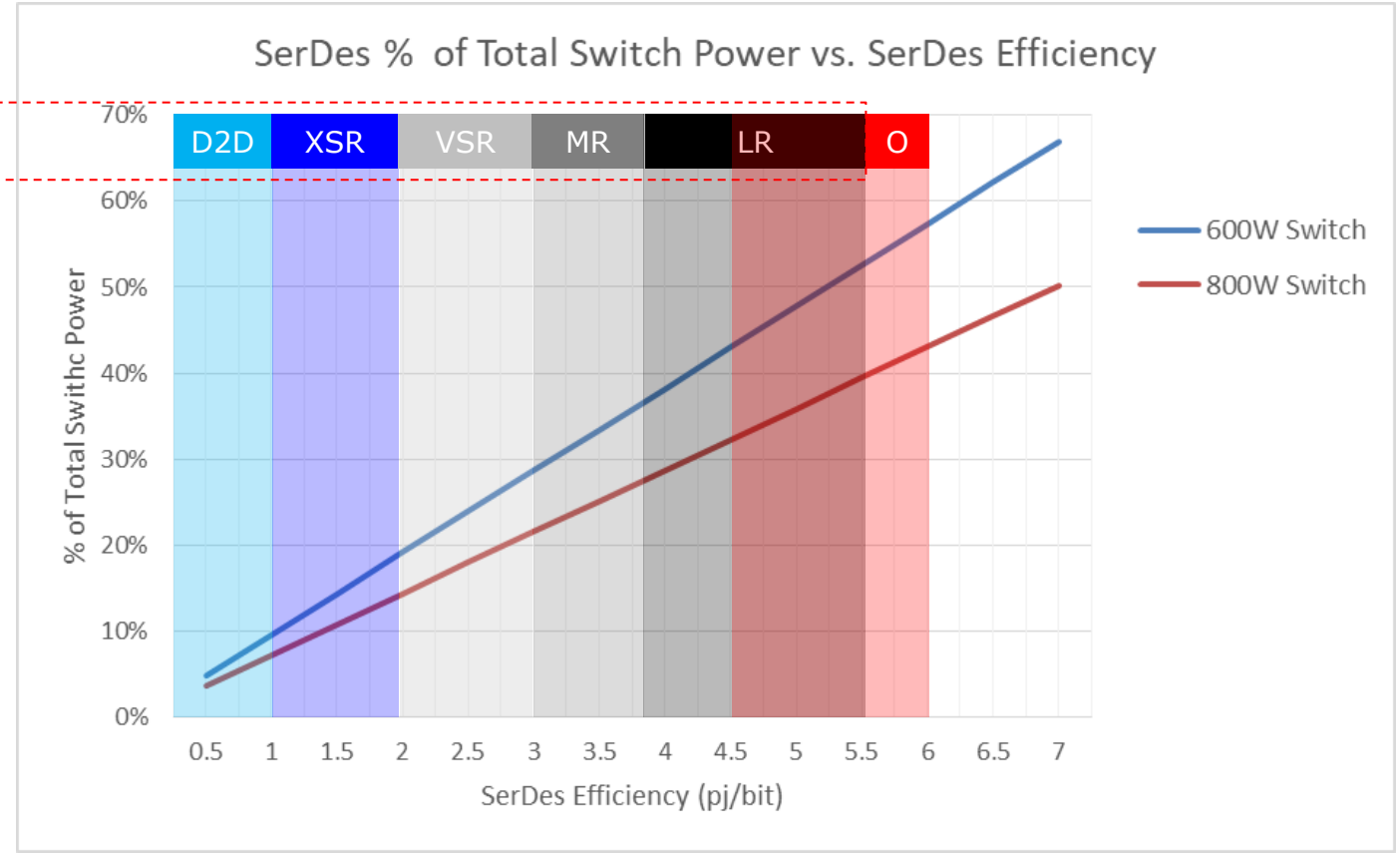
112Gbps PAM4 vs. Technology Node



Curves based on internal estimates & benchmarks not including FEC.
Other non idealities (Xtalk, reflections) will impact this curve.

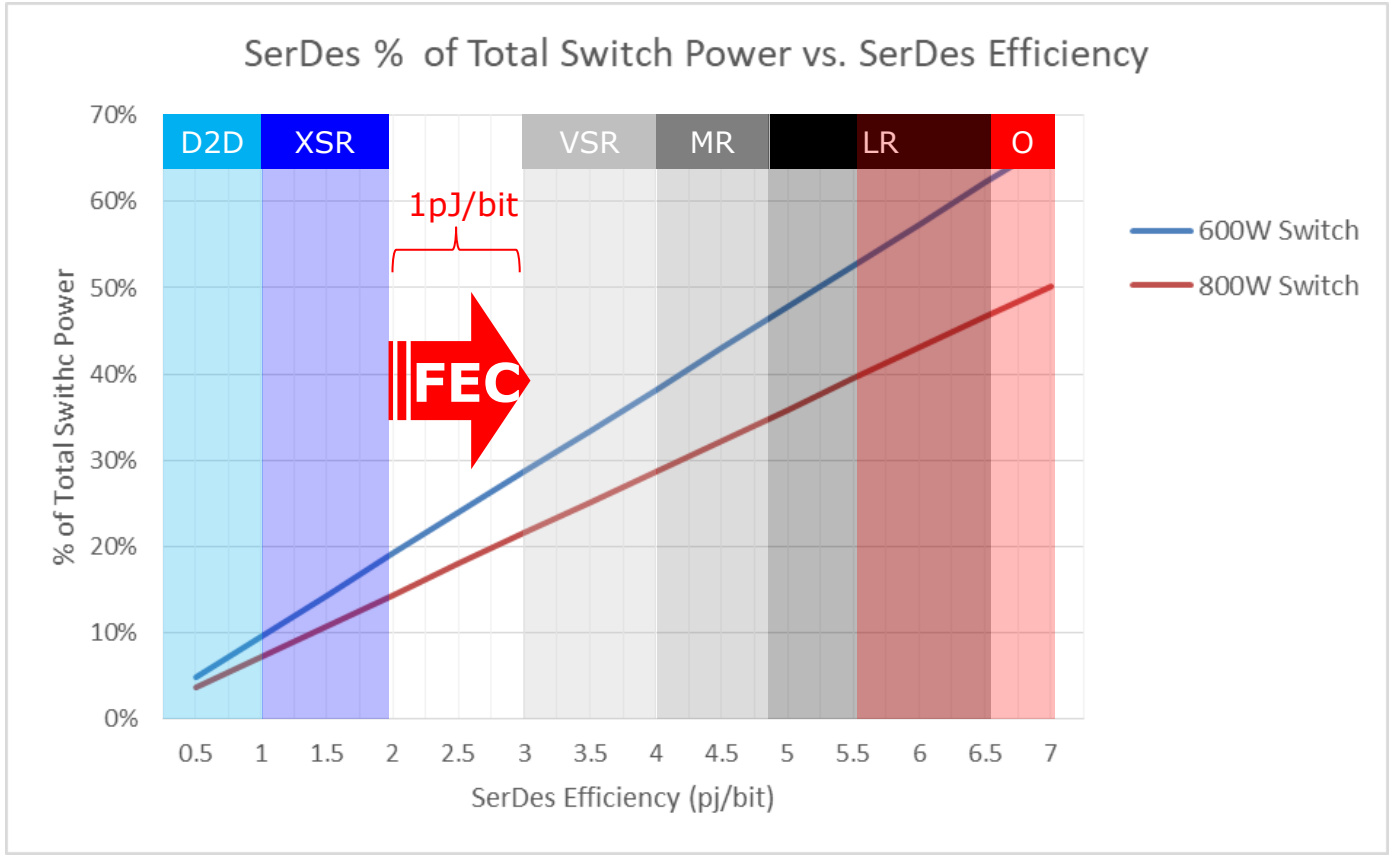
SerDes Power Share: 50T Switch

OIF definition
5nm estimates



SerDes share of total power for a 50T Switch vs. SerDes Energy Efficiency

SerDes Power Share: 50T Switch



SerDes share of total power for a 50T Switch vs. SerDes Energy Efficiency Including KP4 FEC

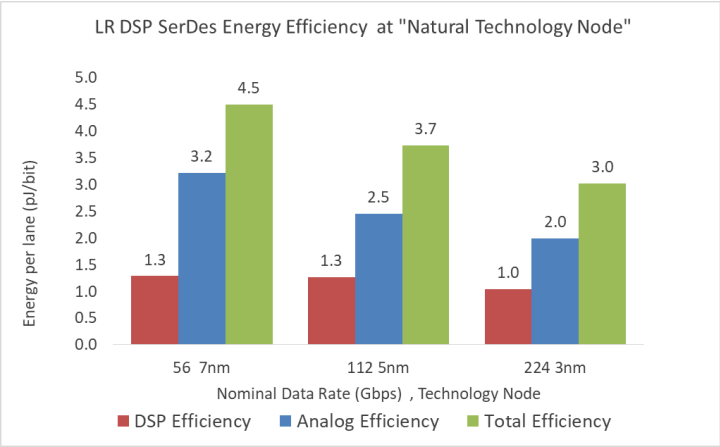
DSP SerDes

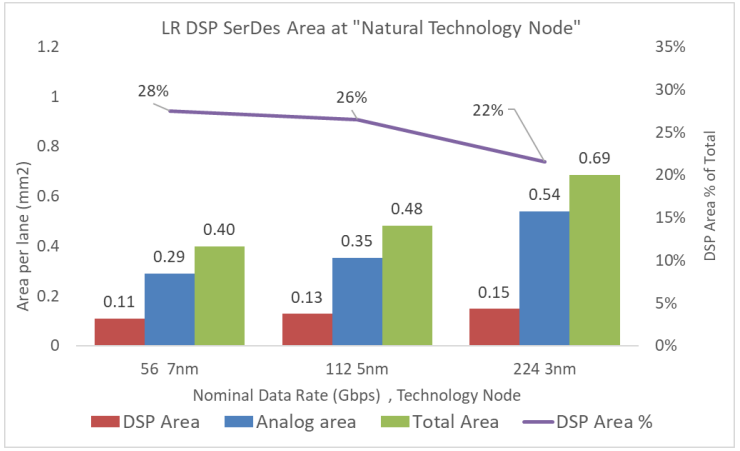
Area and Energy Efficiency Trends

- Targets long electrical channels or short reach optics
- DSP SerDes scales with technology
  - DSP SerDes can improve overall efficiency from 56Gbp to 224Gbps if “natural node” is used
  - “Natural node” : DSP <1/3 total area/power

Parameter	Scaling Factor*	
	7nm→ 5nm	7nm→3nm
AFE Power	0.95	0.9
DSP Power	0.79	0.61
AFE Area	0.95	0.9
DSP Area	0.6	0.36

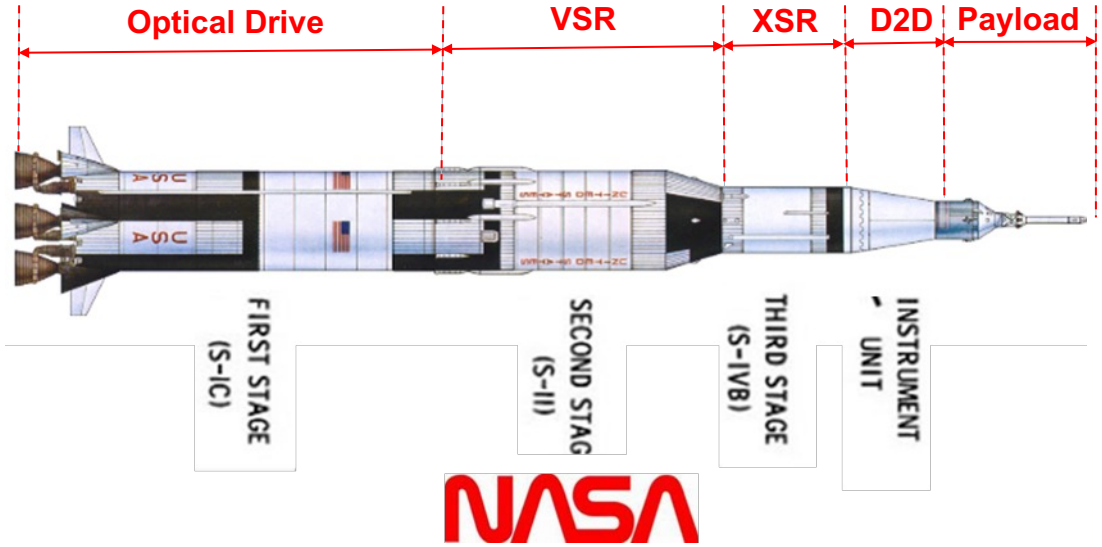
** Technology scaling only, approximate*

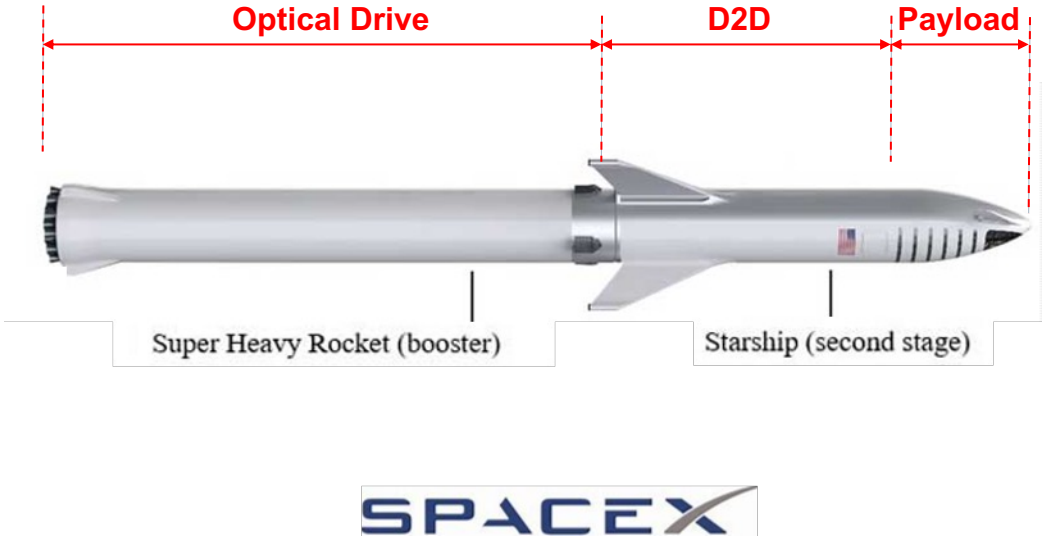


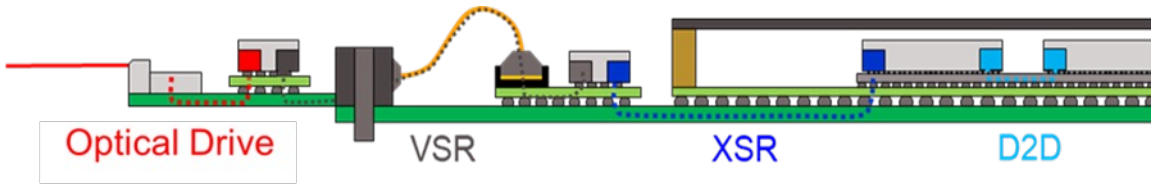


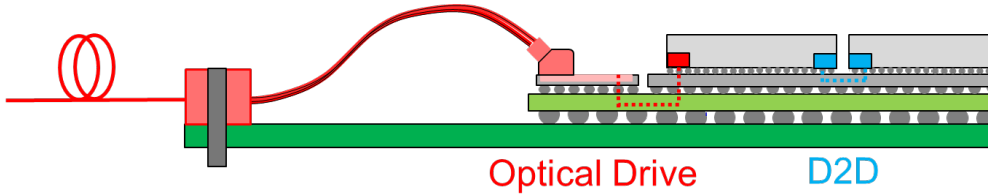
“Natural technology node” :
DSP <1/3 total area/power

Multi Stage vs. Single Stage









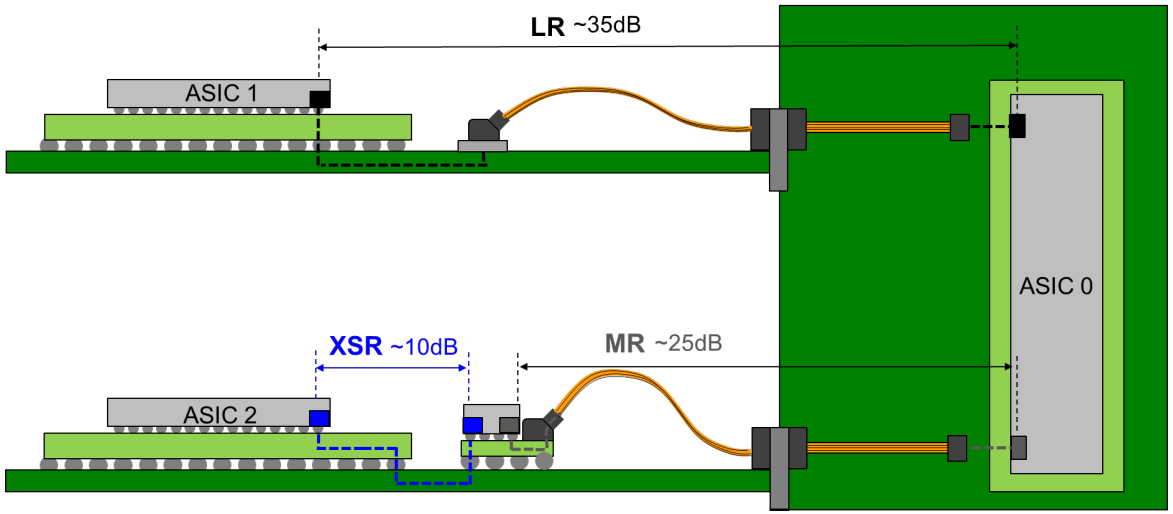
But…..Is single stage always better?

Single vs. 2 Stage

Backplane Example

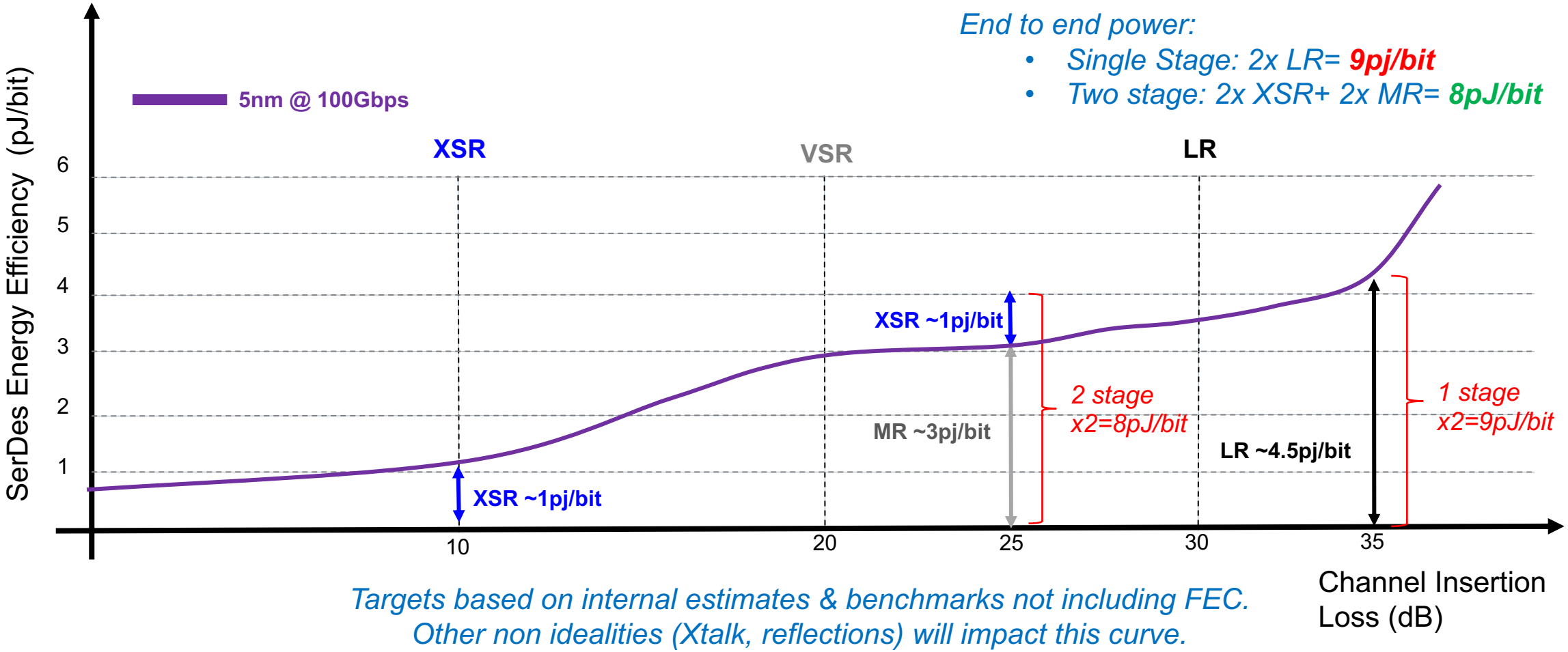
- 100Gbps orthogonal chassis backplane link
- ~35dB bump2bump (B2B) insertion loss (IL)

- Single stage link:
  - LR SerDes in ASIC 1 & 0
  - No repeater
  - Resources: *2x **LR***
- Two stages link:
  - XSR SerDes in ASIC 2, MR in ASIC 0
  - XSR to MR repeater
  - Resources: *2x **XSR** + 2x **MR***



Single vs. 2 Stage

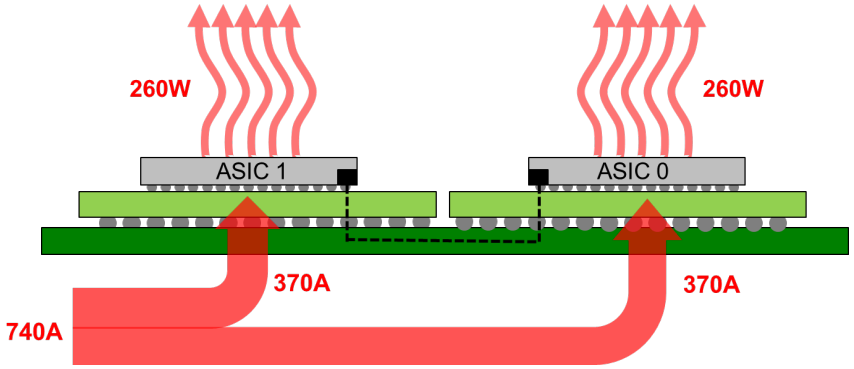
Energy Efficiency

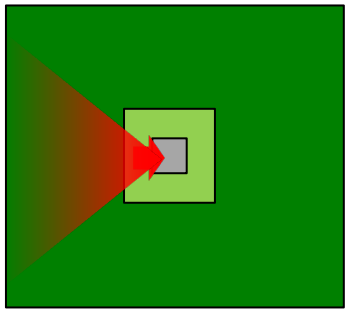


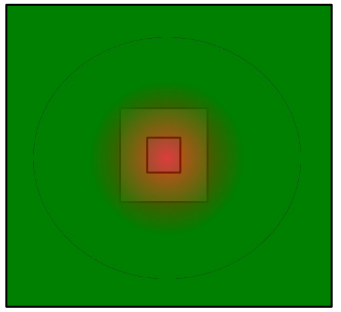
Targets based on internal estimates & benchmarks not including FEC.
Other non idealities (Xtalk, reflections) will impact this curve.

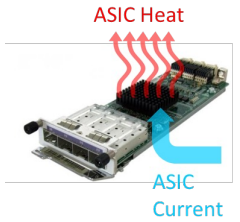
Single vs. 2 Stage

System Considerations



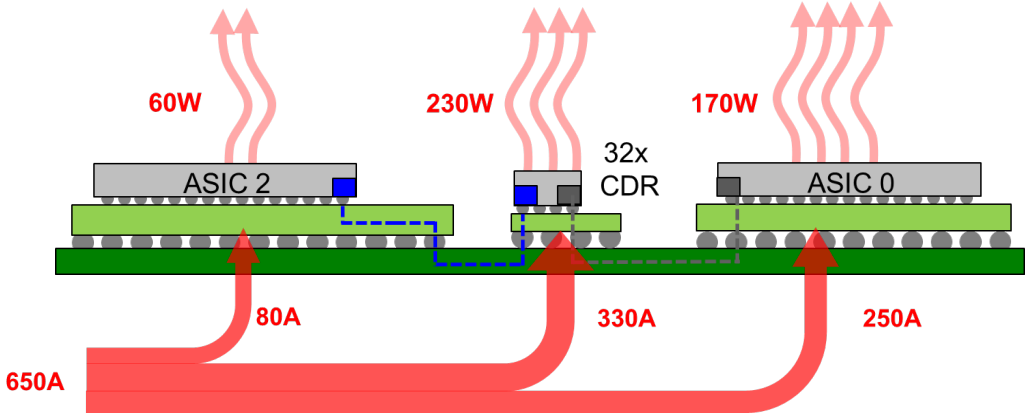


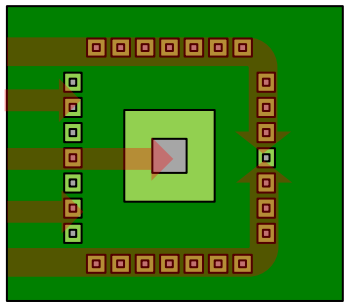


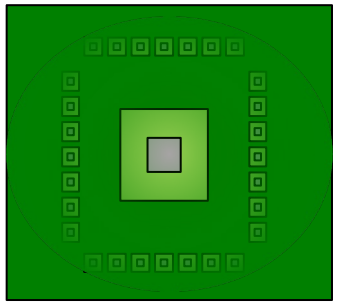


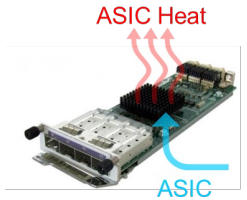
Current Density

Heat



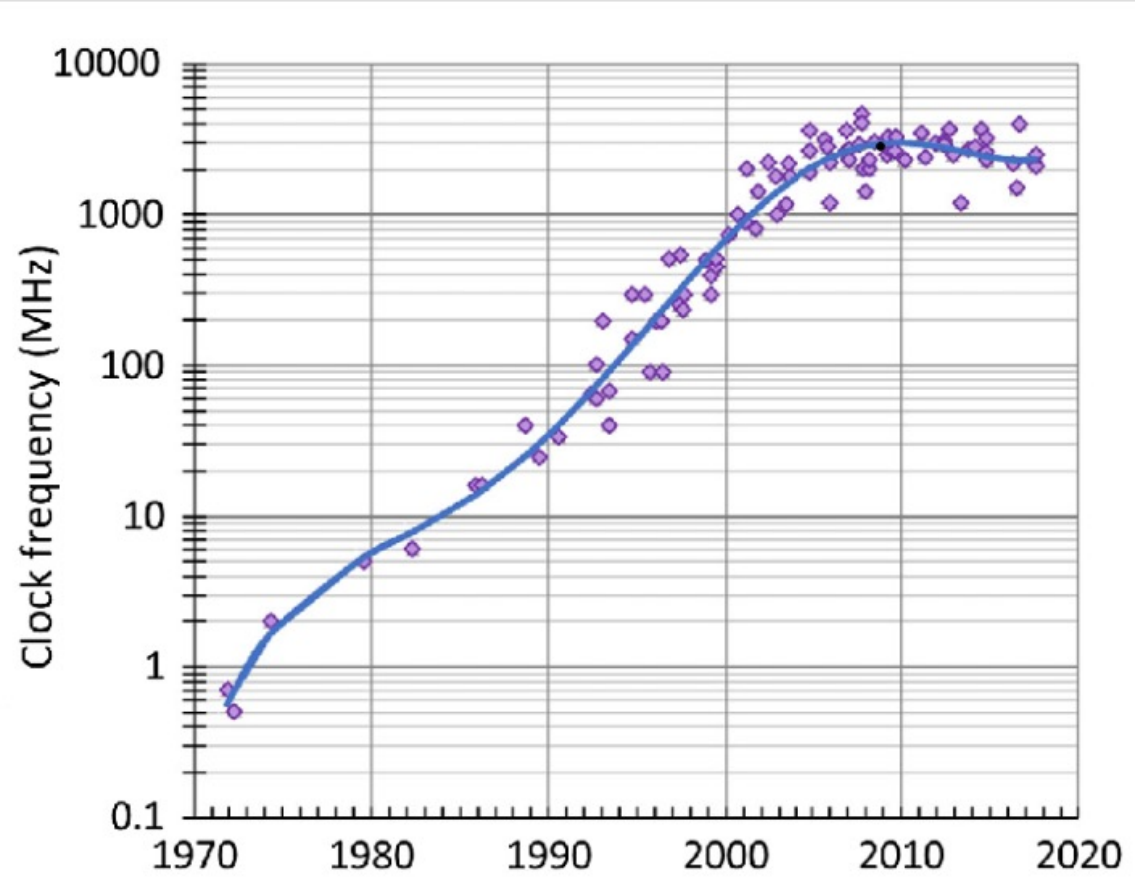






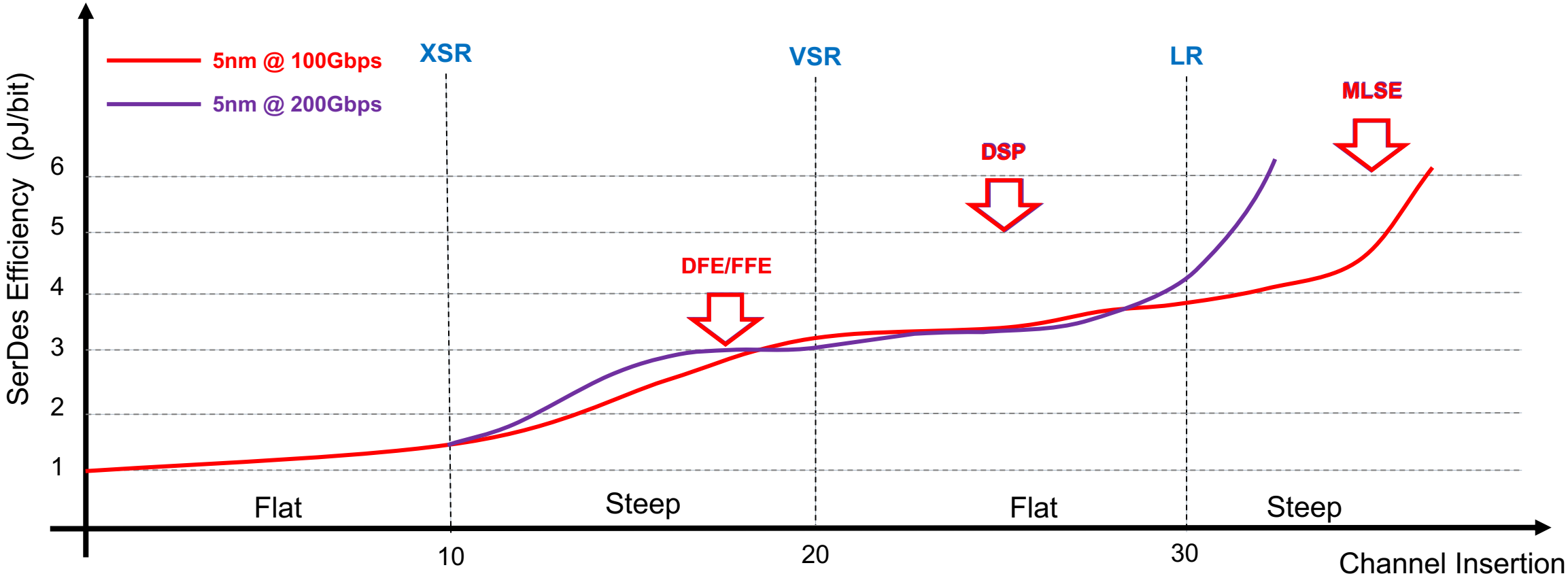
Fast vs. Slow

Anyone Remembers Microprocessors Clock Speed?



Fast vs. Slow

112Gbps vs. 224Gbps Efficiency

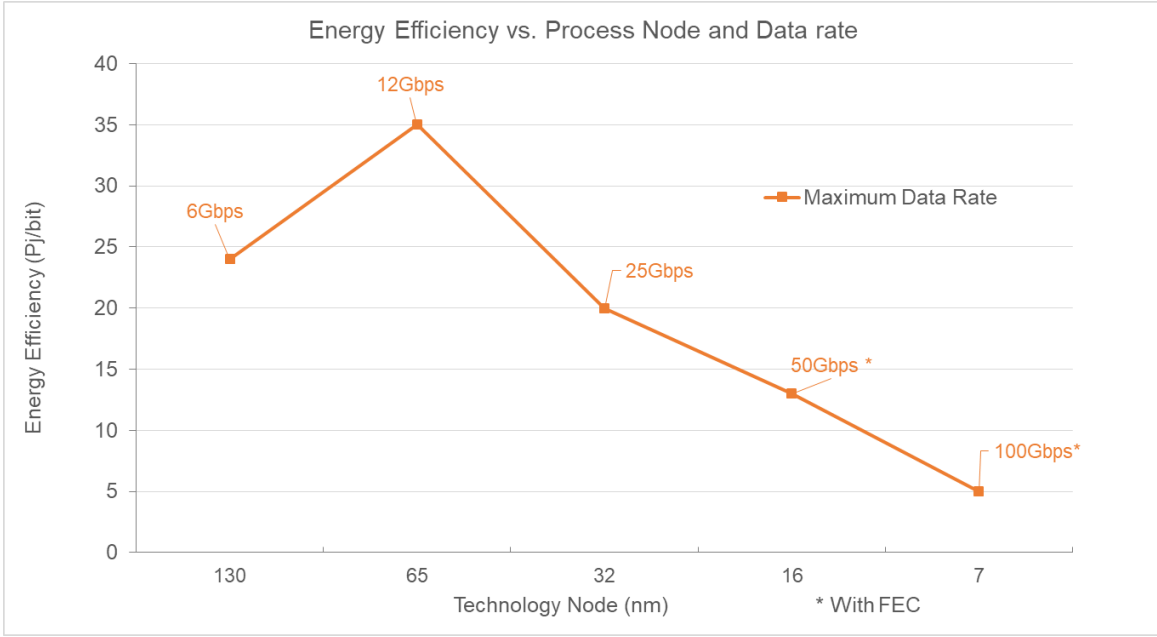


Targets based on internal estimates & benchmarks not including FEC.
Other non idealities (Xtalk, reflections) will impact this curve.

Fast vs. Slow

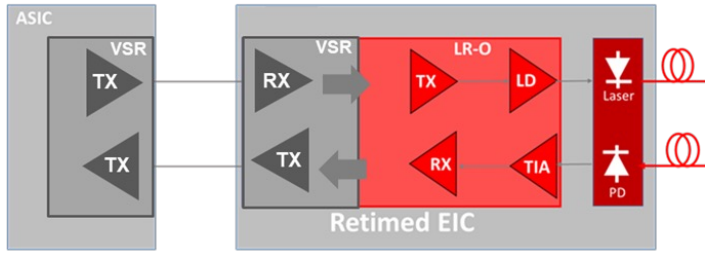
Historical Trends Unpacked

- Ethernet wireline evolved in a connection density constrained environment:
  - Max number of pins, fibers, connectors, balls etc.…
  - Density constraints drove node over node doubling of data rate
- Historical trends could lead us to believe: faster == more efficient
  - Since 10Gbps ~30% node over node
- But built in in this trend:
  - Passive channel improvements
  - CMOS scaling
  - Design and architecture improvements
  - PAM4 efficiency after 50G…
  - ..and higher native BER..
  - …and FEC

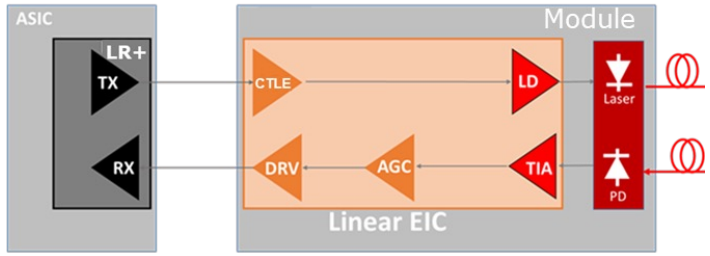


Where Is The Industry Going?

- 800G modules:

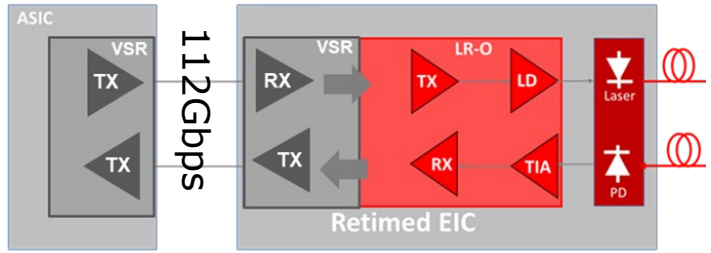


  - Linear: no DSP/ retiming in module
  - More complex SerDes inside ASIC

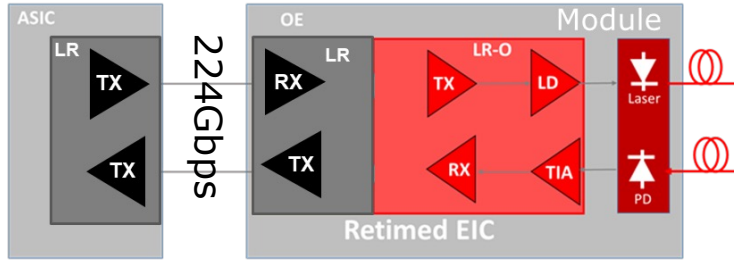


2 stages → 1 stage

- 1.6T retimed modules:



  - Moving to 224Gbps/lane host interface
  - More complex SerDes inside the ASIC

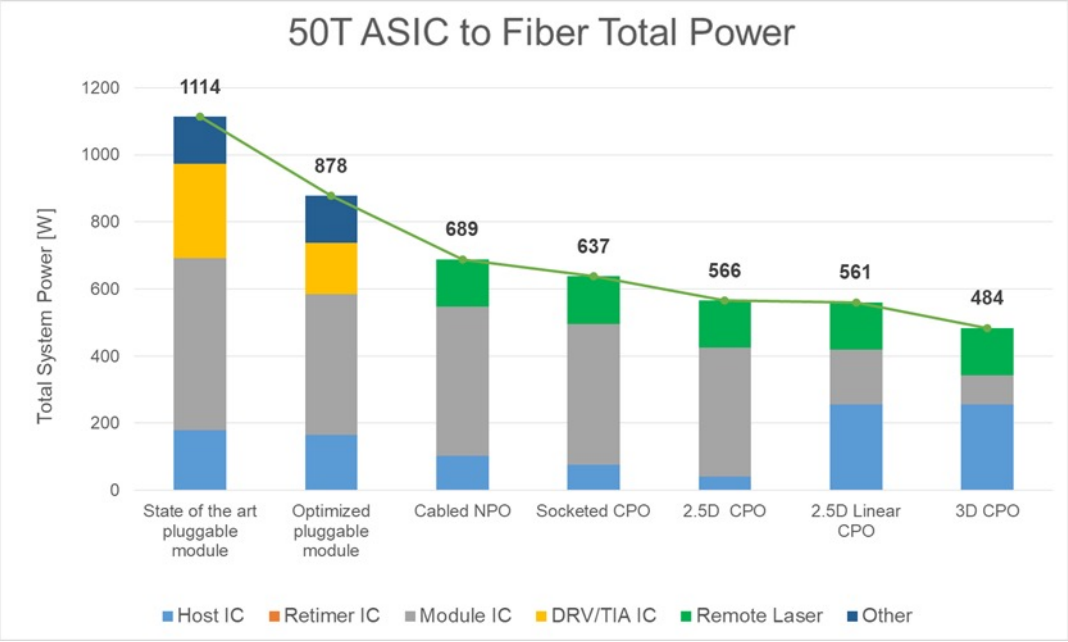


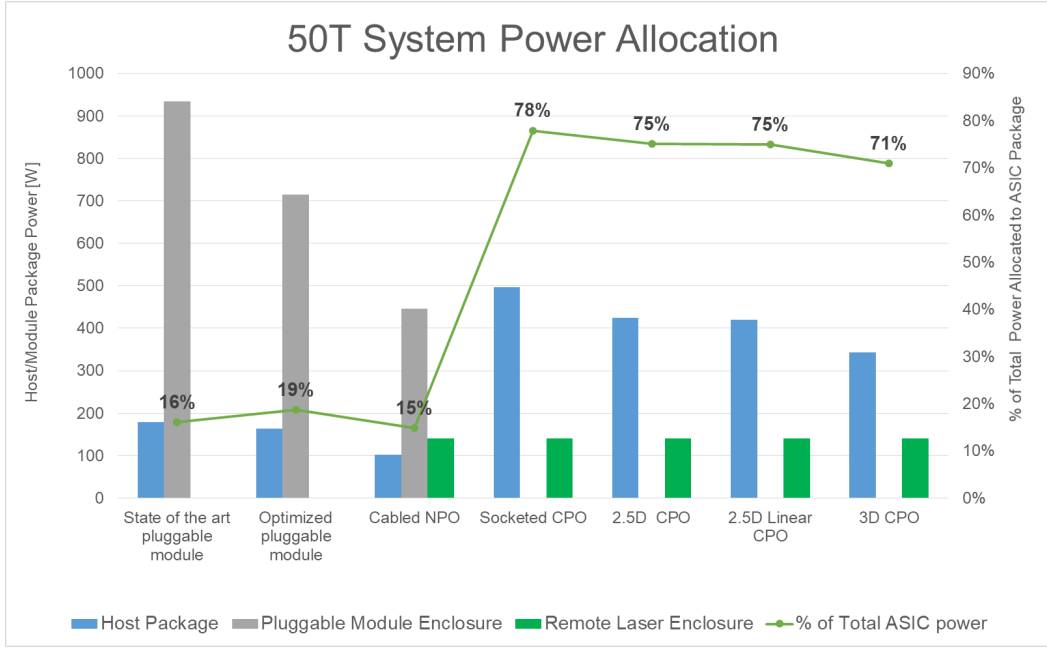
→ faster

Optical Densification

- CPO could reduce system power by ~50% vs. pluggable but…

→ Much more power inside ASIC!!





** 5nm or 7nm is assumed in this analysis*

What About EEE?

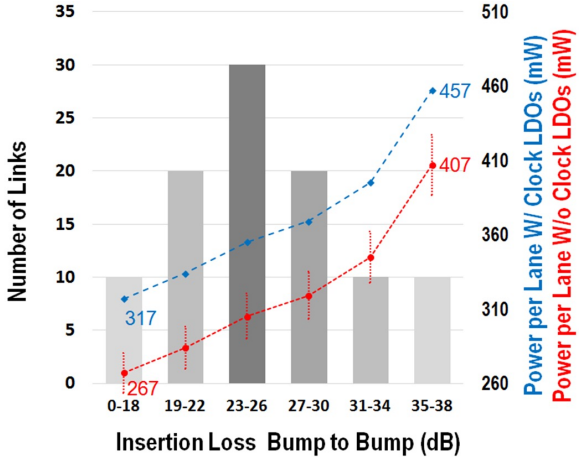
- “Energy Efficient Ethernet”: does it work?

A 12 cylinder Lamborghini that doesn’t idle at the traffic light is….

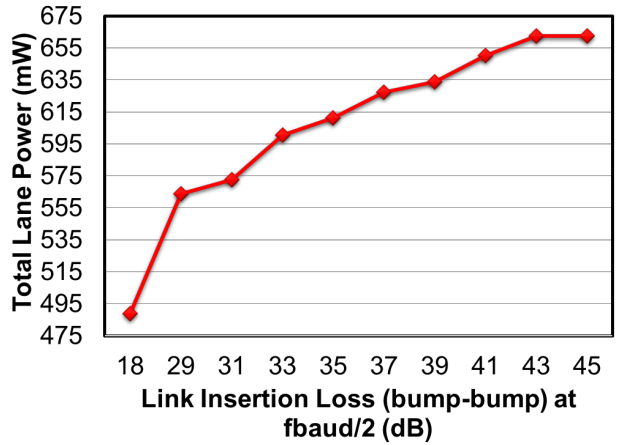
…still a 12 cylinder Lamborghini!!

Power Scaling

- ASICs are usually built with the most performing SerDes available
  - Ports can be used for different applications (front panel, backplane, DAC etc..)
  - Very little flexibility: easy or difficult channels require similar power
- Power scaling can help address efficiency requirements while maintaining flexibility
  - Easy channels require less performance, reduced power consumption



M. LaCroix et al., "A 60Gb/s PAM-4 ADC-DSP Transceiver in 7nm CMOS with SNR-Based Adaptive Power Scaling Achieving 6.9pJ/b at32dB Loss," *ISSCC*, pp. 114-115, 2019.



M. LaCroix et al., "A 116Gb/s DSP-Based Wireline Transceiver in 7nm CMOSAchieving 6pJ/b at 45dB Loss in PAM-4/Duo-PAM-4 and 52dB in PAM-2," ISSCC, pp.132-133, Feb. 2021.

Key Messages

1. SLOW, PARALLEL, SIMPLE *(When possible, but try hard!)*
2. BREAK UP DIFFICULT LINKS *(Always!)*
3. POWER SCALING *(Not just at the traffic light…)*

Thank You