FUJITSU

shaping tomorrow with you

Implementation of DMT transceiver for 400GbE 2km and 10km PMD using 4 wavelengths

Ian Dedic
Fujitsu Semiconductor Europe GmbH

IEEE P802.3bs 400GbE Task Force meeting – San Diego, July 2014

Supporters

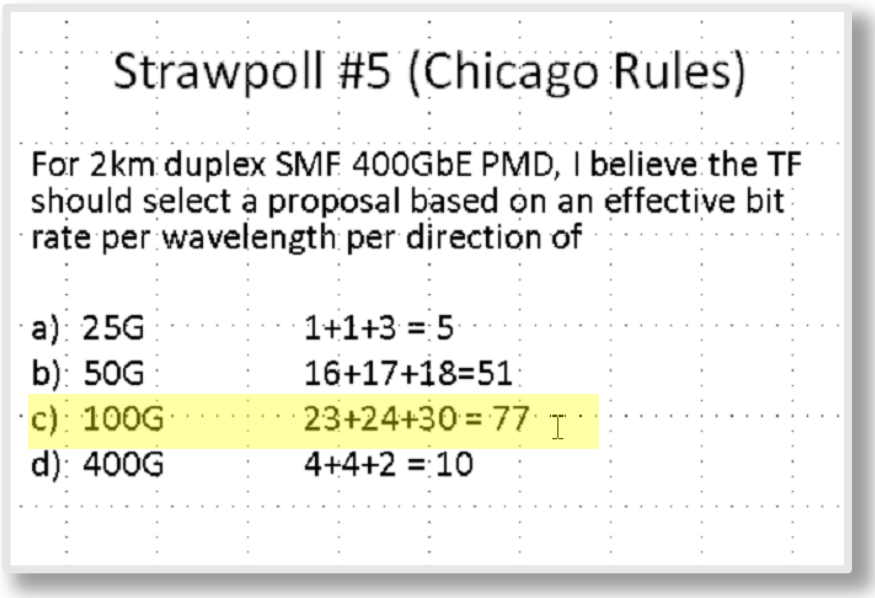
- David Lewis, JDSU
- Sacha Corbeil, JDSU
- Beck Mason, JDSU
- Moon S Park, OE Solutions
- Tom McDermott, Fujitsu Network Communications
- Rolf Steiner, Agilent
- Bernd Nebendahl, Agilent

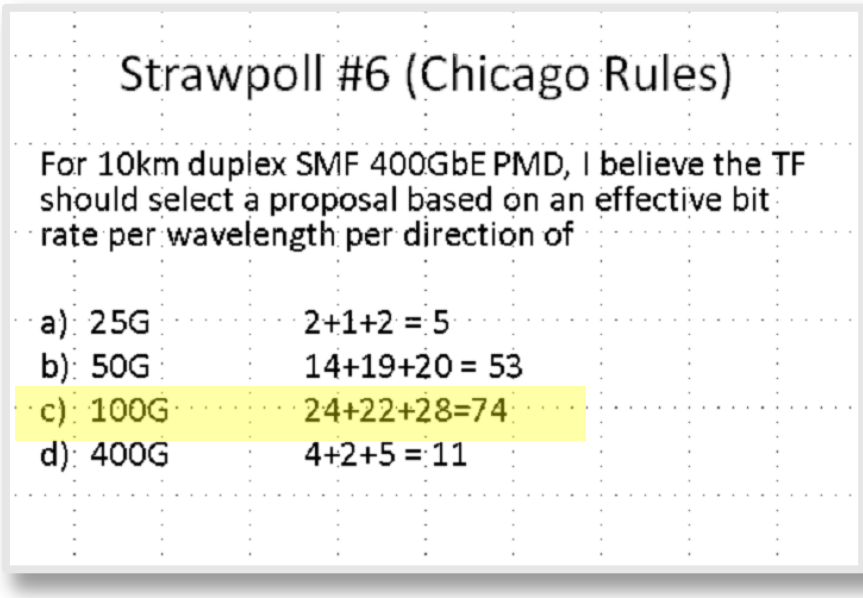
400GbE with 4-Lambda

FUJITSU

- Motivation?
  - Reduction of optical component count needed to reduce cost
  - Allows WDM without making lasers too difficult/expensive
  - Cost-effective building block needed for next generation 400G and beyond

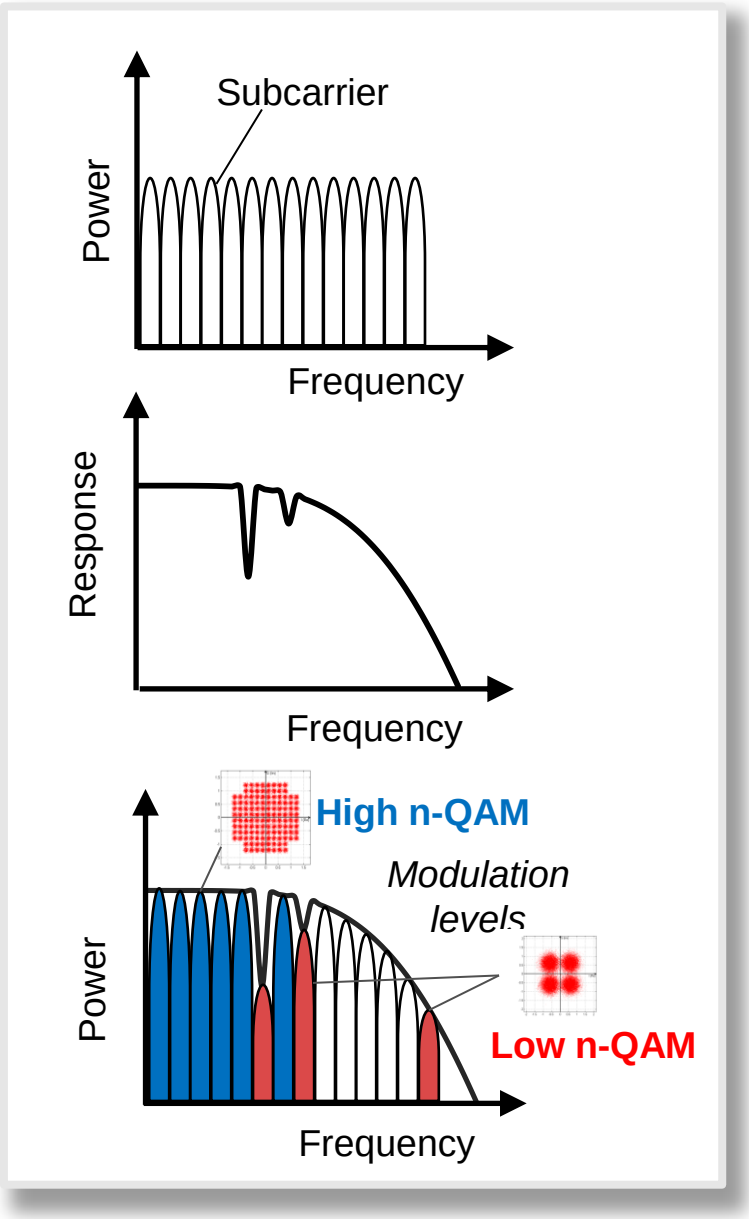
- Straw Polls from Norfolk IEEE 802.3bs interim meeting indicated strong interest in 100Gbps/lambda solutions for 400GbE





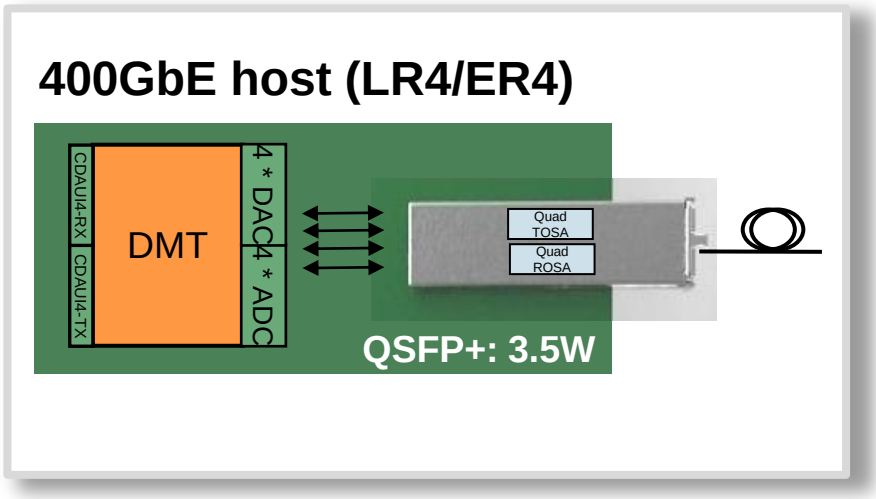
Discrete Multi-Tone Quick Overview

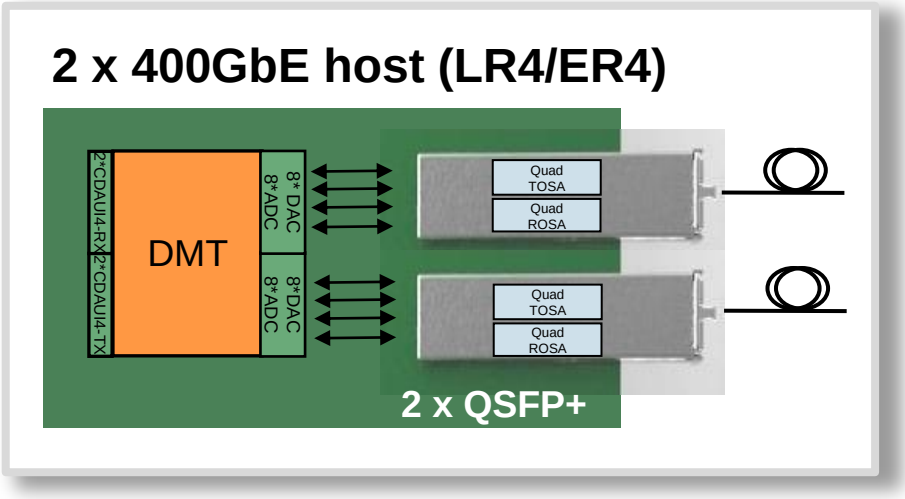
- DMT uses a series of uniformly separately subcarriers (e.g. 256) to transmit data, where each subcarrier uses QAM-n modulation
- Level of modulation is adjusted based on the available SNR within each subcarrier to maximize the number of bits per symbol within that subcarrier
- Transmitter and Receiver communicate to determine optimal settings for transmission
- Flexible modulation on each of the subcarriers allows DMT to compensate for link impairments and achieve the best use of the available signal channel bandwidth and SNR
- Frequency domain calculations with efficient IFFT & FFT implementations
- Transmit data is assembled in the frequency domain before passing through an IFFT and high speed DAC
- Receive DMT data is captured using a high-speed ADC, converted back into the frequency domain via FFT where subcarrier amplitude and phase are mapped into received data



400G DMT Transceivers

- Can support optical PMDs for 500m, 2km, 10km
- Can potentially support flexible line rates
- 4 x 1-ch 100G transceivers or dedicated 4-ch 400G transceiver
- Host-based transceiver provides opportunity for delivering 400G in a QSFP+ form factor
  - Optics only inside QSFP+ module, with 4*100G DMT electrical interface
  - 4*116Gbps electrical interface from chip to module through connector
  - DMT can cope with imperfections of pluggable module connector



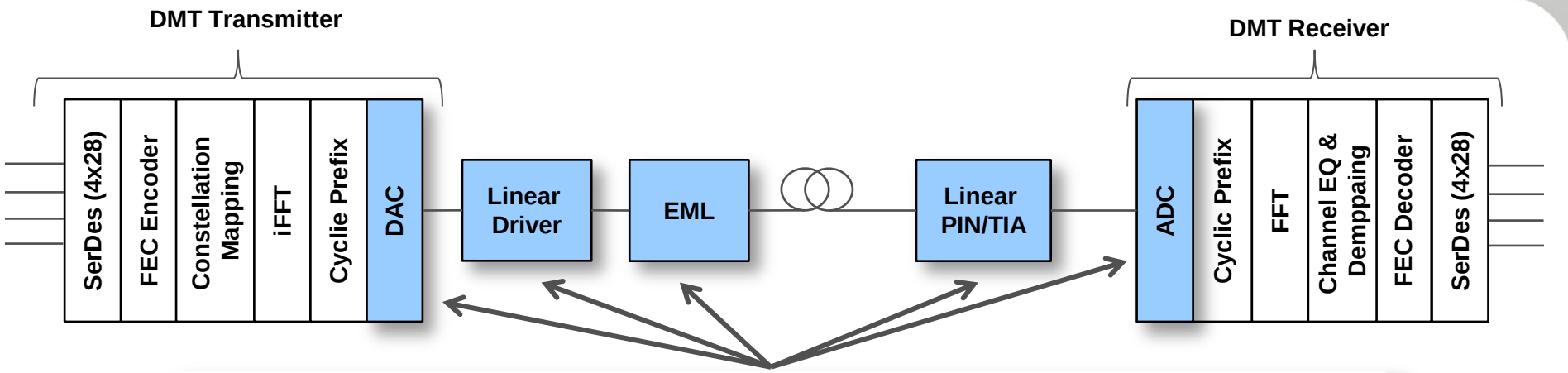


System Requirements

FUJITSU

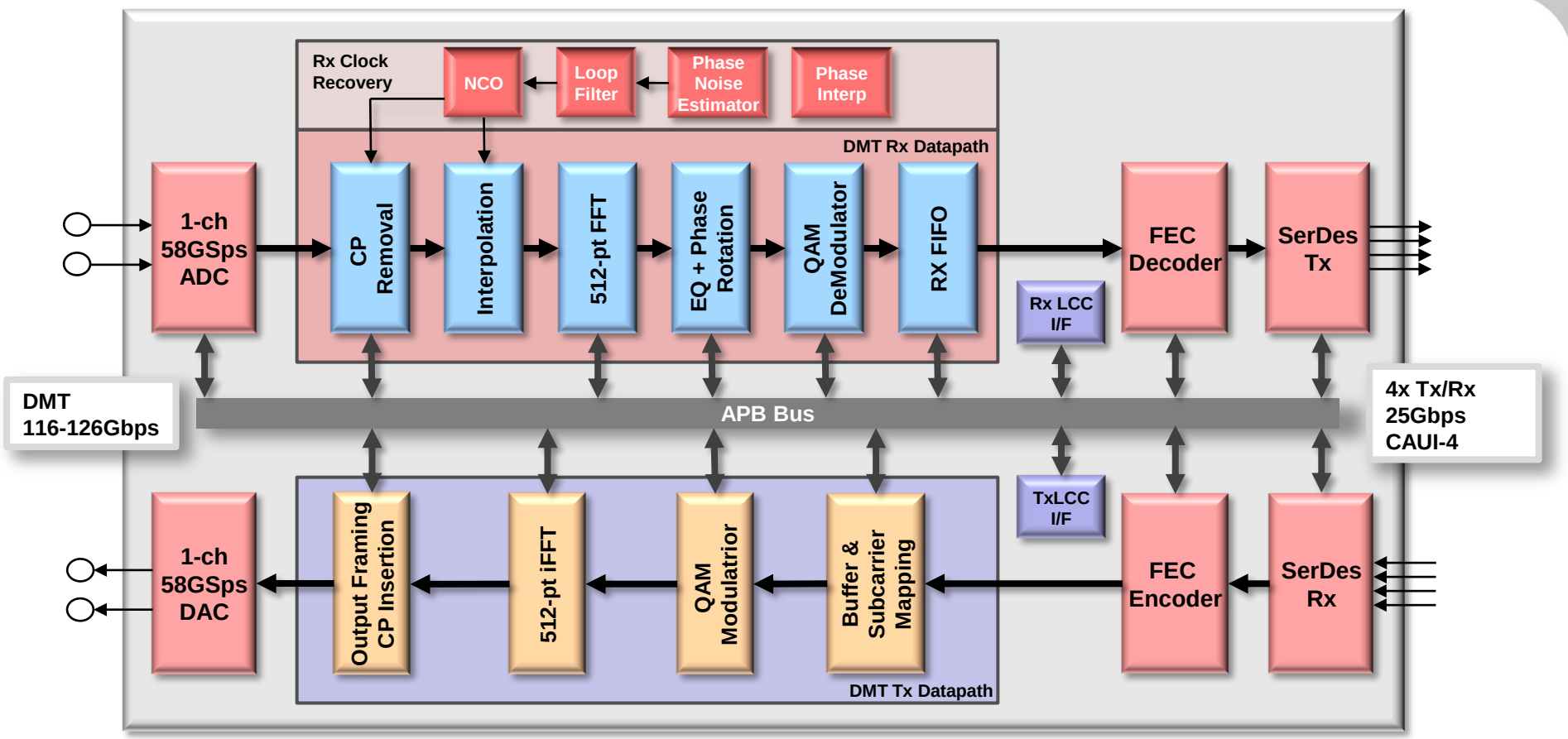
Block		DMT Requirement (100Gbps/lambda)
ADC	Sample Rate	58GSa/s
	ENOB	> 5
	Bandwidth	18GHz
DAC	Sample Rate	58GSa/s
	ENOB	> 5
	Bandwidth	16GHz
SerDes	4-lane 28Gbps	CEI-28G-VSR
FEC	Coding Gain	~8dB
DMT DSP	Gate Count	~5M
Linear TIA		< 25G class
Linear Driver		< 25G class
Laser (DML/EML)		< 25G class

Bandwidth Requirements (DMT)



- 25G-class optics (maybe even lower bandwidth) can be used
- Existing ADC & DAC (28nm) can already generate 170Gbps electrical back-to-back transmission (@ BER 1e-03) over “bandwidth limited components”
  - Increases in bandwidth in next process node will add more margin for ROSA and TOSA
- Packaging & PCB losses @ 16GHz-20GHz manageable
- Bandwidth is never free
  - Lower bandwidth requirements will reduce overall system difficulty & cost
  - Higher digital power in DMT engine compensated for by power savings in optics
    - In reality this power saving may be bigger than added DMT power cost...

100G DMT Transceiver IC Overview



- Single chip integrated CMOS transceiver for 100Gbps/lambda
  - Low cost, low loss 10mm x 10mm SHDBU package for QSFP28 module integration
  - 3.5W in 28nm (silicon measurement, synthesis and layout results) – OK for CFP4
  - **<1.8W in 14nm (estimated from "100G-type" benchmarks) – OK for QSFP28**

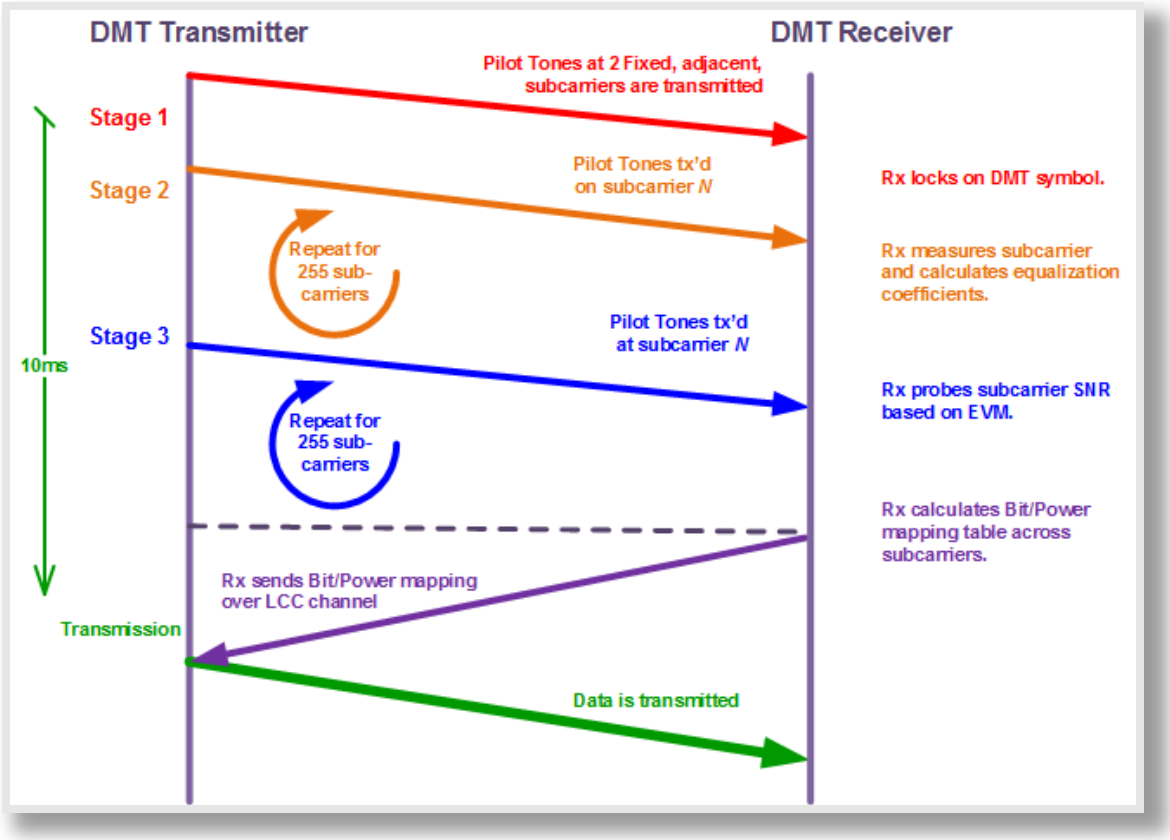
FFT/IFFT and CP Implementation

FUJITSU

- FFT & IFFT can be efficiently realized in dedicated hardware
  - Mixed-radix FFT with hardcoded twiddle factors (directly synthesized, no actual multipliers)
  - Core clock frequency Fsamp/128 (450MHz at 58Gs/s)
    - Optimises power efficiency of digital (parallel processing of 128 samples)
- 2 selectable FFT sizes, 64SC (128 point FFT) and 256SC (512 point FFT)
  - 64SC lower power/latency/capacity (SNR margin), 256SC higher power/latency/capacity
    - Pilot tones and cyclic prefix add less overhead with 256SC
    - 256SC for longer reach or to allow lower performance (cheaper) optics
  - Basic FFT is 128 points in 1 clock cycle – used for 64SC
    - Mixed-radix (radix-2 and radix-4) to minimise power and simplify twiddle factors
  - 256SC FFT adds a final radix-4 stage every 4 clock cycles (small power increase)
- Selectable cyclic prefix (CP) length
  - Longer CP on poorer channels (longer impulse response, more reflections)
  - Shorter CP on better (shorter?) channels ➔more margin ➔allows use of 64SC mode
- FFT size and CP length could be fixed or negotiated during startup/probing
- Power and area are much smaller than calculated from “number of MACs”
  - FFT + IFFT power is <300mW in 28nm for 256SC
  - DMT TX+RX (datapath+clock recovery) power is <900mW (~25% of total device power)

DMT Link Negotiation & Communication

- Known location subcarriers are used to create a Link Communication Channel (LCC using DC FFT bin) and for frame/bit alignment
- DMT requires initial negotiation between the TX and RX on startup
- TX sends known pilot tones (fixed constellation and bit sequence) for probing channel SNR via EVM across all 256 subcarriers
- RX analyzes SNR and exchanges back bit allocations tables with TX over LCC channel
- Target link negotiation time around 10ms from cold

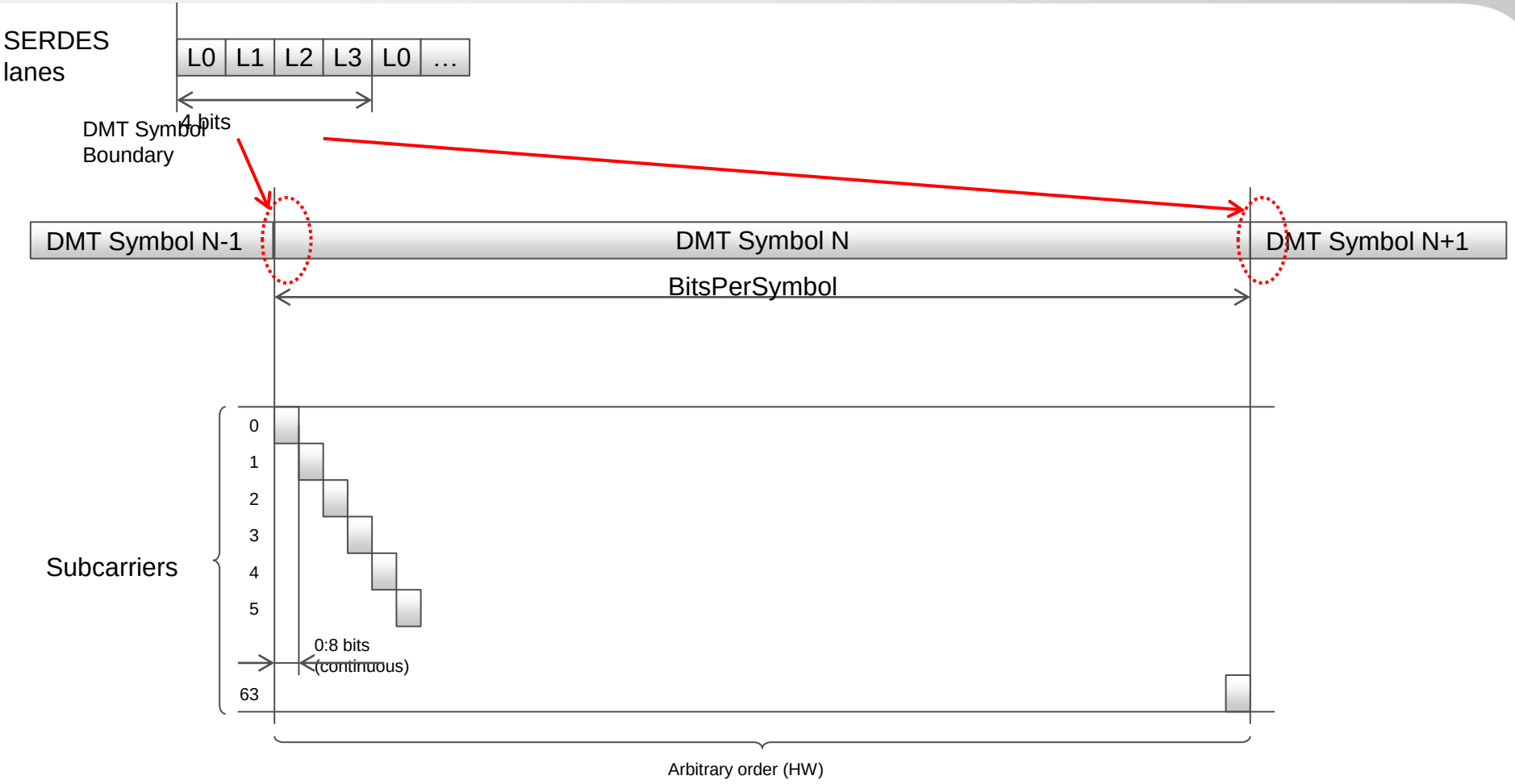


Bit allocation scheme

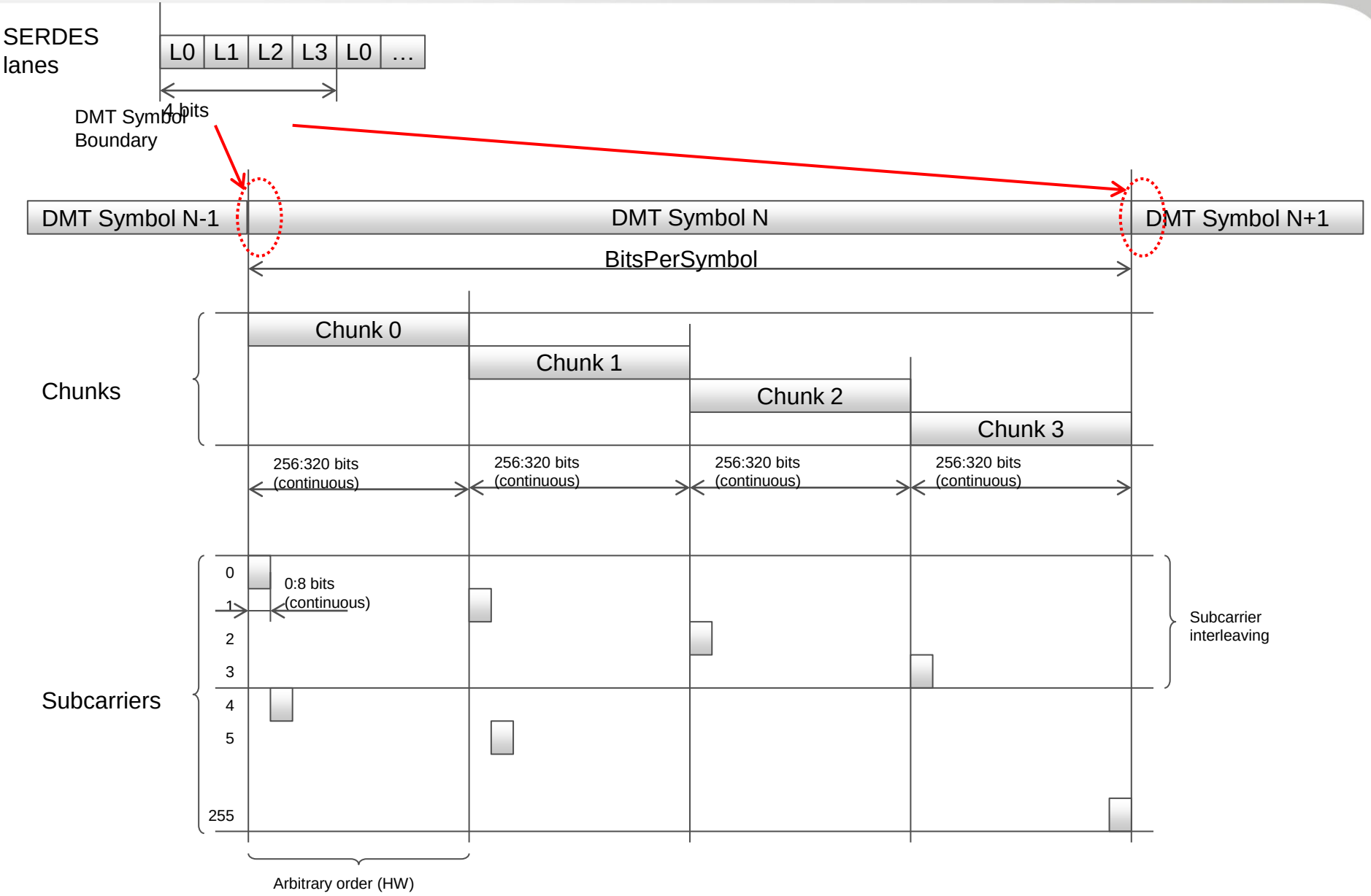
FUJITSU

- Bits arriving at DMT Tx input are allocated into DMT symbols
  - To achieve a target bit rate each DMT symbol has to carry a specified number of bits
  - The number of bits per DMT symbol is divisible by 4 to maintain SERDES lane ordering
  - SERDES lanes are bit-interleaved at the input of DMT Tx
- Bits allocated into a 256-subcarrier DMT symbol are further divided into 4 chunks
  - Avoids buffering of the whole DMT symbol – reduces latency and power consumption
    - Performance optimization, not an inherent requirement of a DMT link
  - Each chunk is a continuous sequence of input bits
  - Chunks may have unequal sizes
    - Size range 256:320 bits or 256:296 bits
  - To minimize inequality of chunk sizes due to channel roll-off, chunks are mapped to interleaved subcarriers
  - In 64-sub-carrier mode there is only 1 chunk per DMT symbol, no interleaving, no overlaps.
- Bits allocated into a chunk are distributed among 64 subcarriers
  - Each subcarrier carries 0 to 8 bits
  - Each group of 0-8 bits is a continuous sequence of input bits
  - Hardware implementation does not impose restrictions on subcarrier ordering
    - Allocation algorithm allocates chunks to subcarriers in ascending order

Bit mapping diagram, SC=64



Bit mapping diagram, SC=256



Number of bits per DMT symbol

FUJITSU

Number of Sub-Carriers, SC [Samples]	Cyclic Prefix Size, CP [Samples]	Bits/DMT Symbol, BitsPerDMT [bits]	Nominal Chunk Size, BitsPerChunkNom [bits]	TxIn-RxOut Alignment, BitAlign [bits]
64	2	260	260	4
64	4	264	264	8
64	8	272	272	16
64	12	280	280	24
64	16	288	288	32
256	4	1032	258	8
256	8	1040	260	16
256	16	1056	264	32
256	32	1088	272	64
256	48	1120	280	96
256	64	1152	288	128

BitsPerDMT = 2*(2*SC+CP)

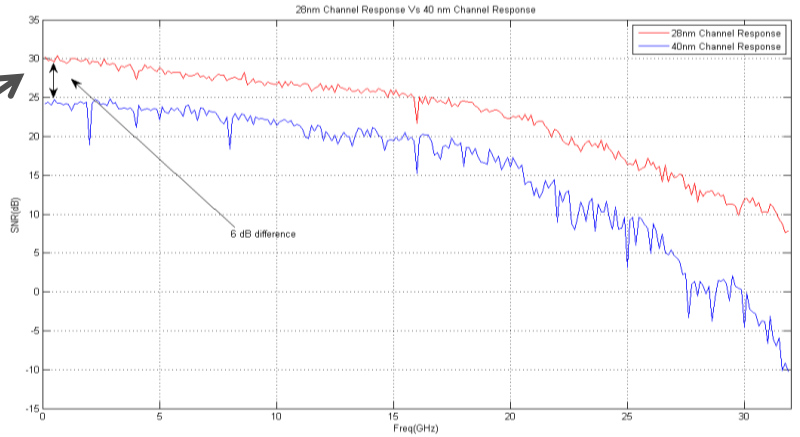
BitsPerChunkNom =

BitsPerDMT, in SC=64 mode

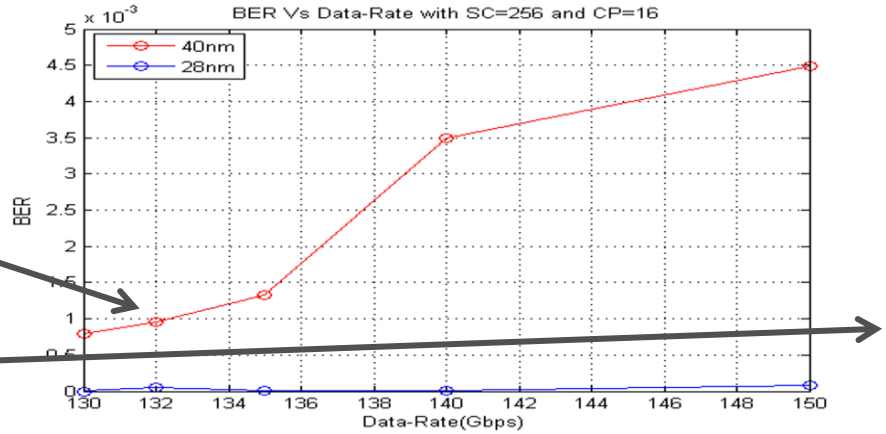
BitsPerDMT/4, in SC=256 mode

DMT Electrical B2B Measurements

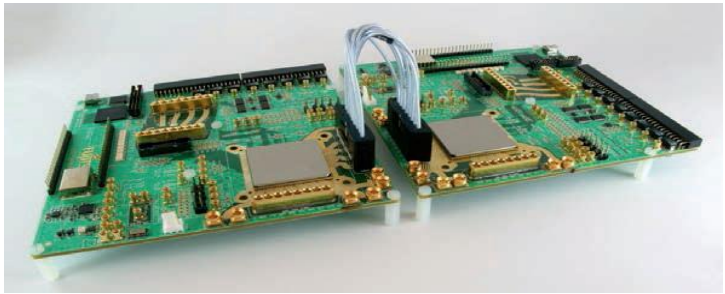
- Electrical B2B measurements made using existing ADC & DAC silicon
  - Channel probing shows ~6dB SNR improvement between 40nm and 28nm
  - Excess bitrate (above 116Gbps at 1e-3 BER) gives indication of relative margin in the system for other component losses



- 40nm Silicon Measurements
  - **132Gbps** throughput
  - 14% bitrate margin
- 28nm Silicon Measurement
  - **170Gbps** throughput
  - 46% bitrate margin



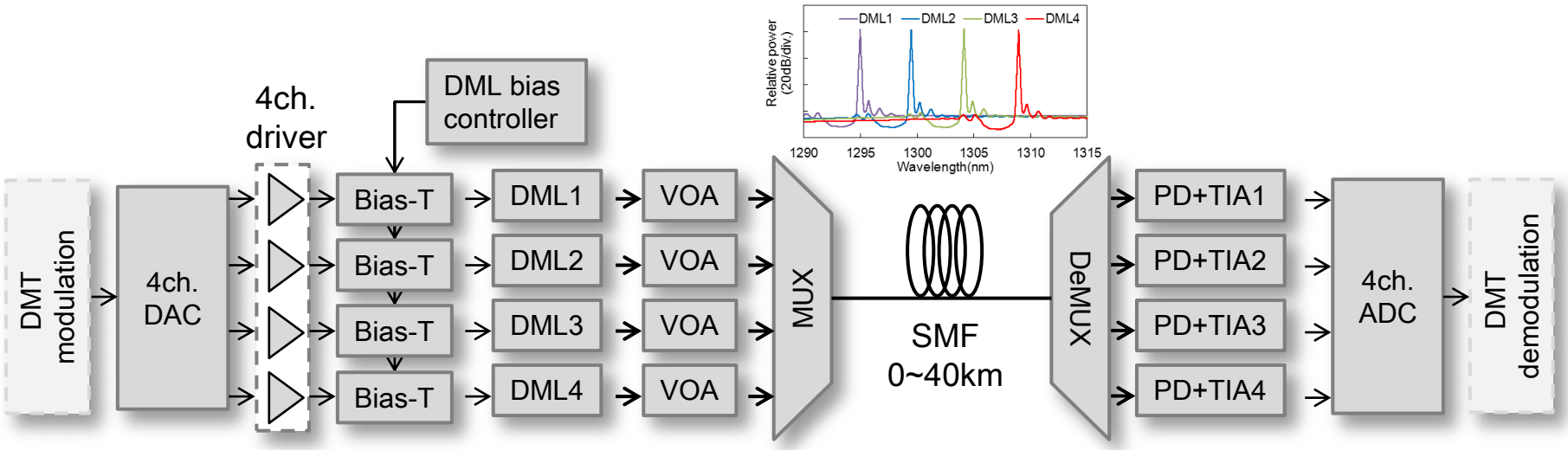
- Initial optical experiments using 28nm converters presented, work in progress

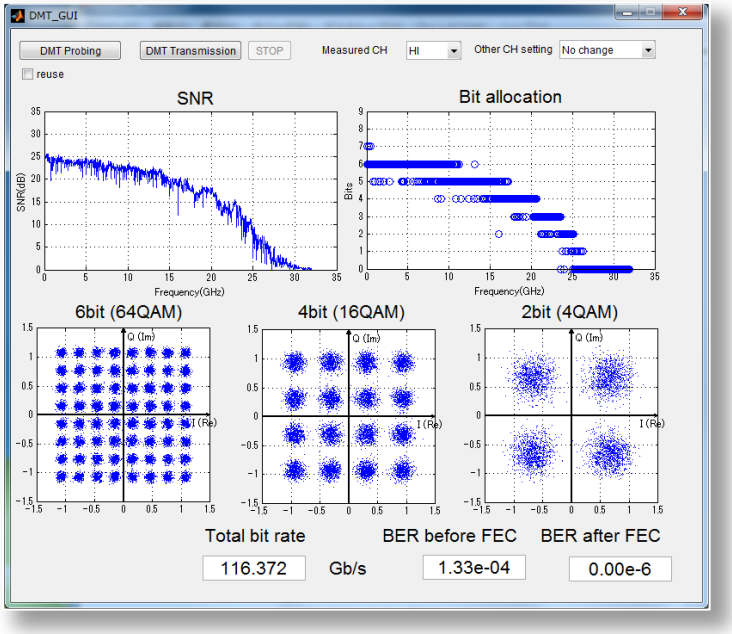


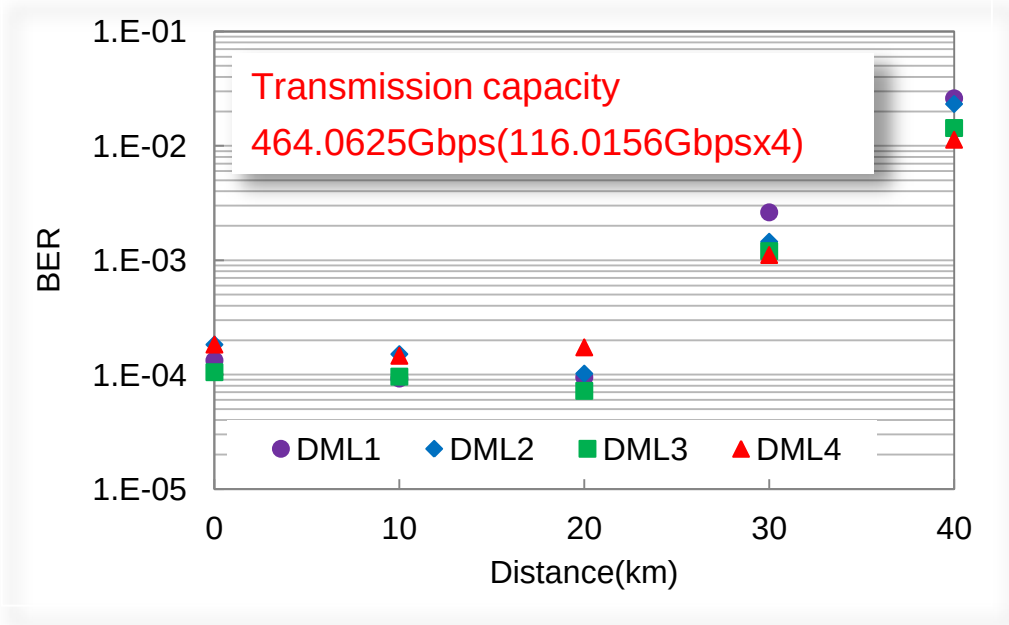
Plots show 256 subcarriers, Cyclic Prefix of 16
Bit rate measured at FEC threshold of 1e-3

Optical 400GbE DMT Experiment (40nm)

FUJITSU

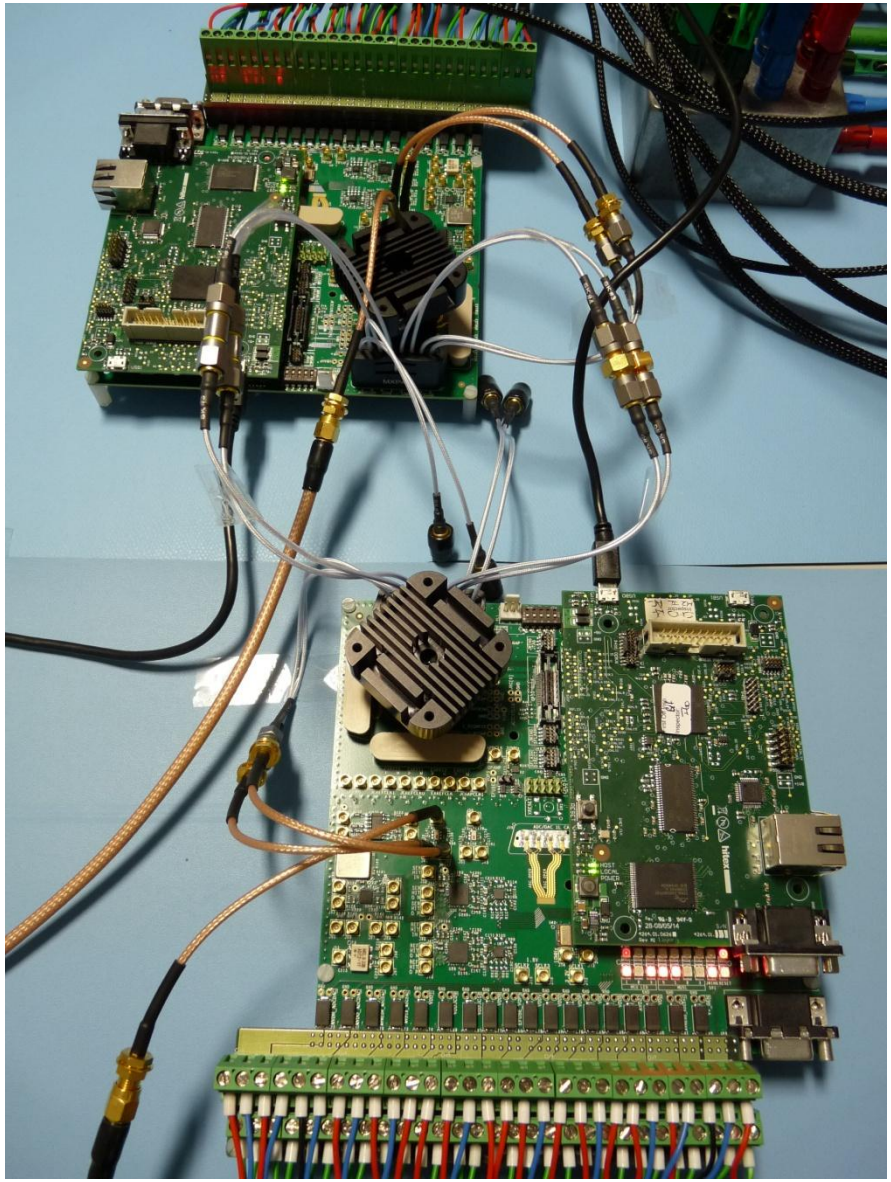


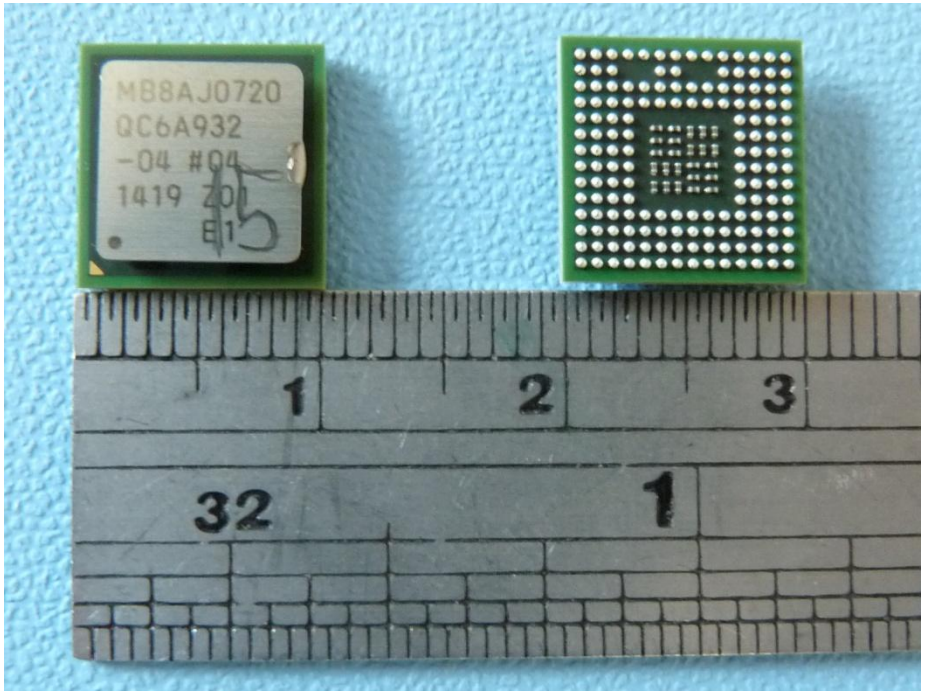




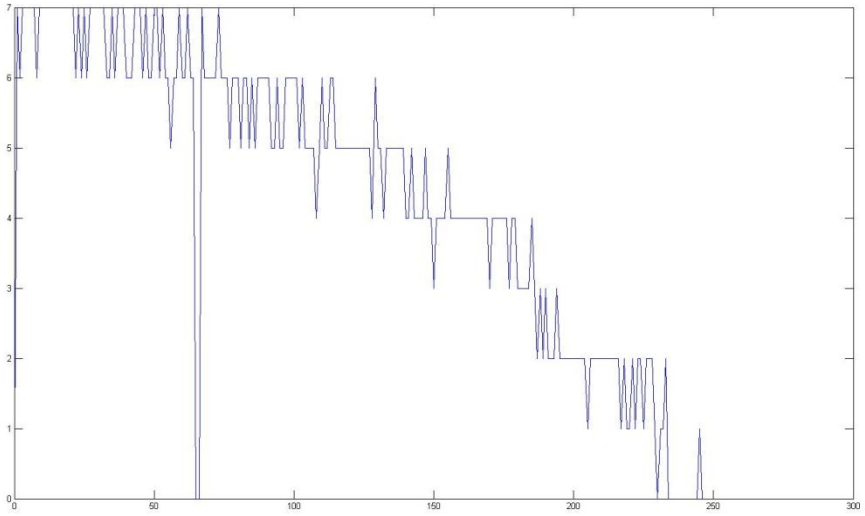
100G DMT Transceiver – not a simulation...

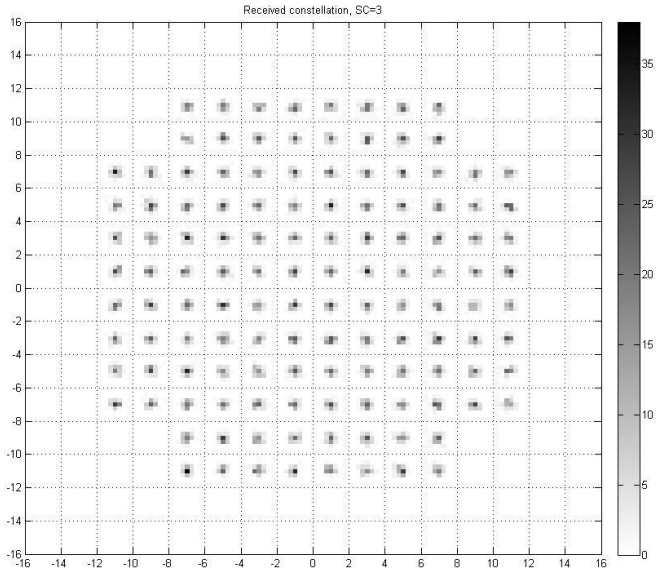
FUJITSU

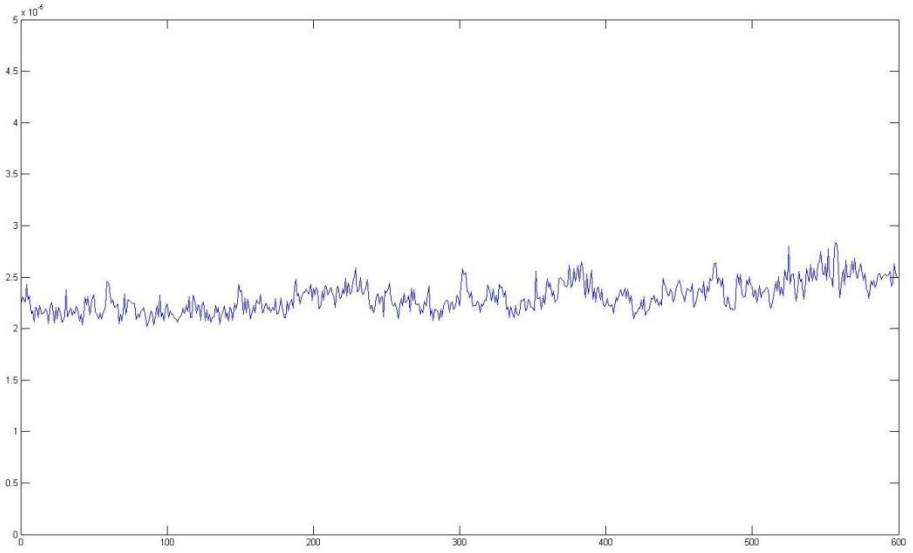


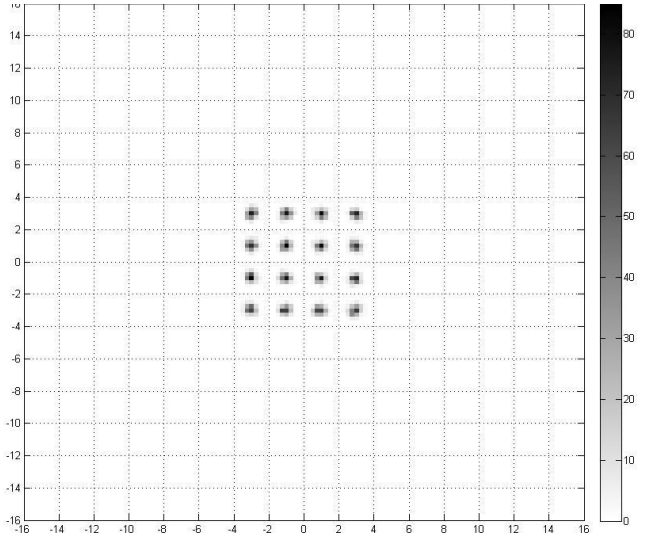


100G DMT Transceiver – measured results









Summary

FUJITSU

- 400GbE with 4 lambda is readily achievable through use of DMT modulation and ADC/DSP/DAC architectures
  - CMOS technology can offer performance and cost advantages
  - Allows host-based solution with only low-power optics in pluggable module
  - Enables low-cost single-lambda 100G and roadmap to 4-lambda 1TbE (no standards!)
- Implementation of DSP solutions is not a concern
  - Silicon issues already solved for coherent long-haul
  - Can help decrease complexity/cost/power in the optics
  - Link negotiation/retraining is really a non-issue (well-known and solved in DSL)
- Bandwidth requirements need to be carefully considered
  - Bandwidth isn't free, there are always tradeoffs
    - Lower bandwidth generally equals lower cost and power
  - Focus should not only be on just the DSP silicon, but all components