Towards an objective for 400 Gb/s for DCI applications

Markus Weber, Tom Williams - Acacia
Gary Nicholl, Mark Nowell - Cisco
Tad Hofmeister - Google
Ilya Lyubomirsky - Inphi
Jeffrey Maki - Juniper
Rich Baca, Brad Booth, Mark Filer - Microsoft

Supporters

- Your name here

Topics

- Background
- Market Need
- Technical Feasibility and Leverage
  - 400G Extender Sublayer
  - PCS
  - FEC
  - Optics
- Economic Feasibility
- Summary
- Proposed Objective
- Backup

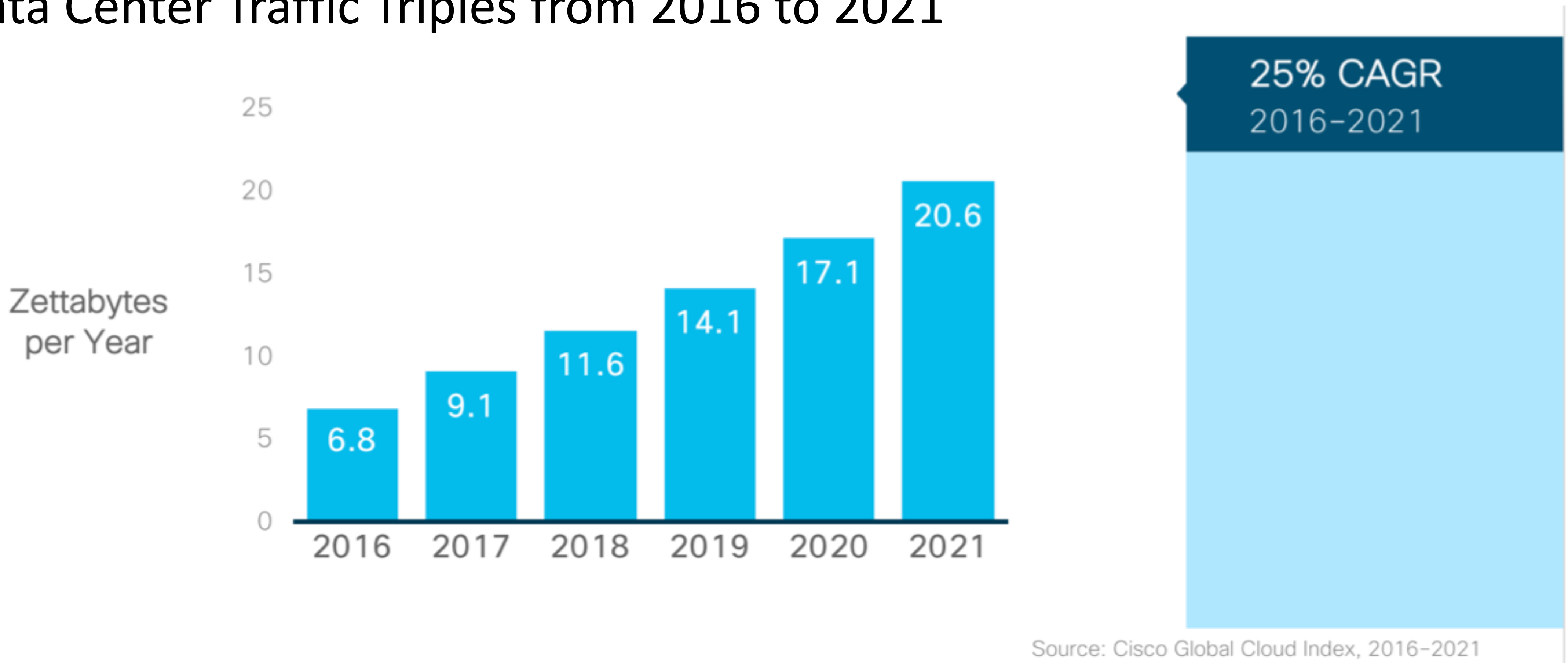
Background

- There is interest within the industry in defining new Ethernet DWDM PHYs with the ability to run over a single-channel (wavelength) port on a DWDM multichannel optical system.
- At the Chicago meeting (March, 2018) an objective was adopted to support 100 Gb/s operation on a single wavelength capable of at least 80km over a DWDM system, and initially targeting the Cable/MSO market.
- This presentation proposes to adopt a similar objective to support 400 Gb/s operation on a single wavelength capable of at least 80km over a DWDM system, and initially targeting the DCI market.

Market Need

Global Data Center Traffic Growth

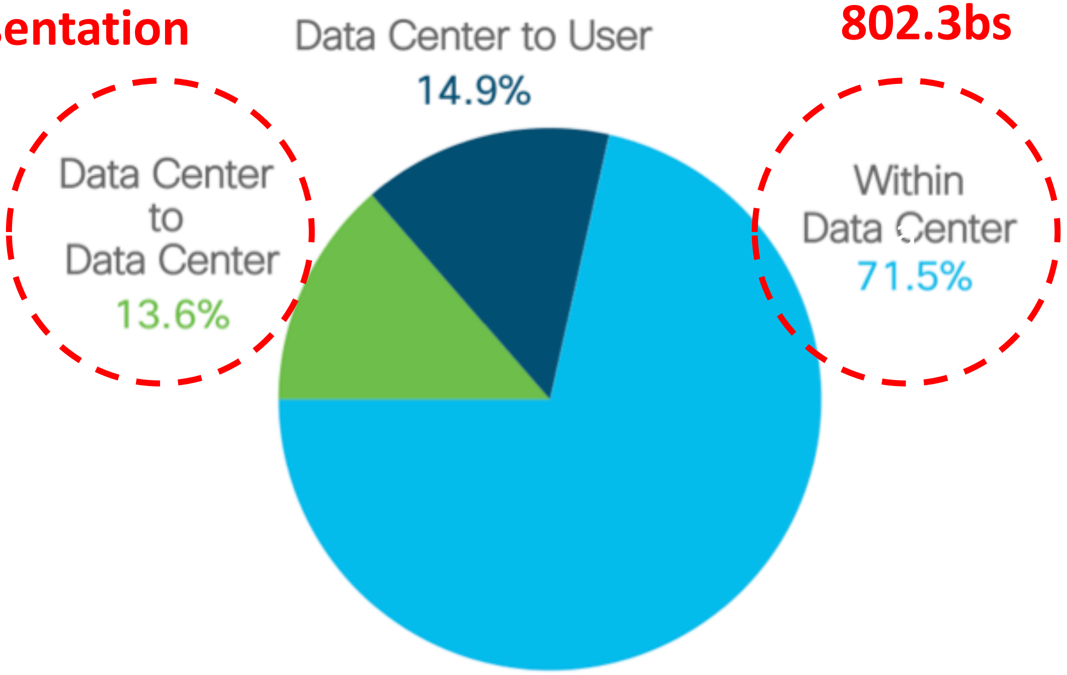
Data Center Traffic Triples from 2016 to 2021



Global Data Center Traffic by Destination, 2021

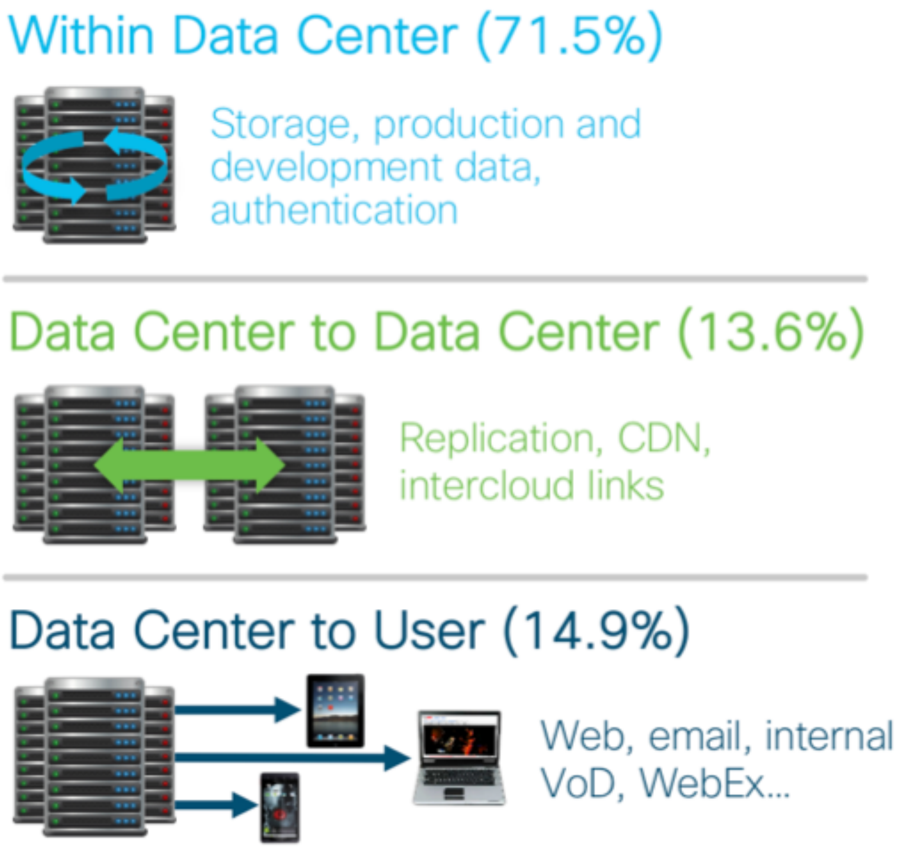
Target of this presentation

Addressed by 802.3bs



Total East-West Traffic Will Be 85%

(Rack-local traffic would add another slice twice the size of "Within Data Center")



Source: Cisco Global Cloud Index, 2016–2021

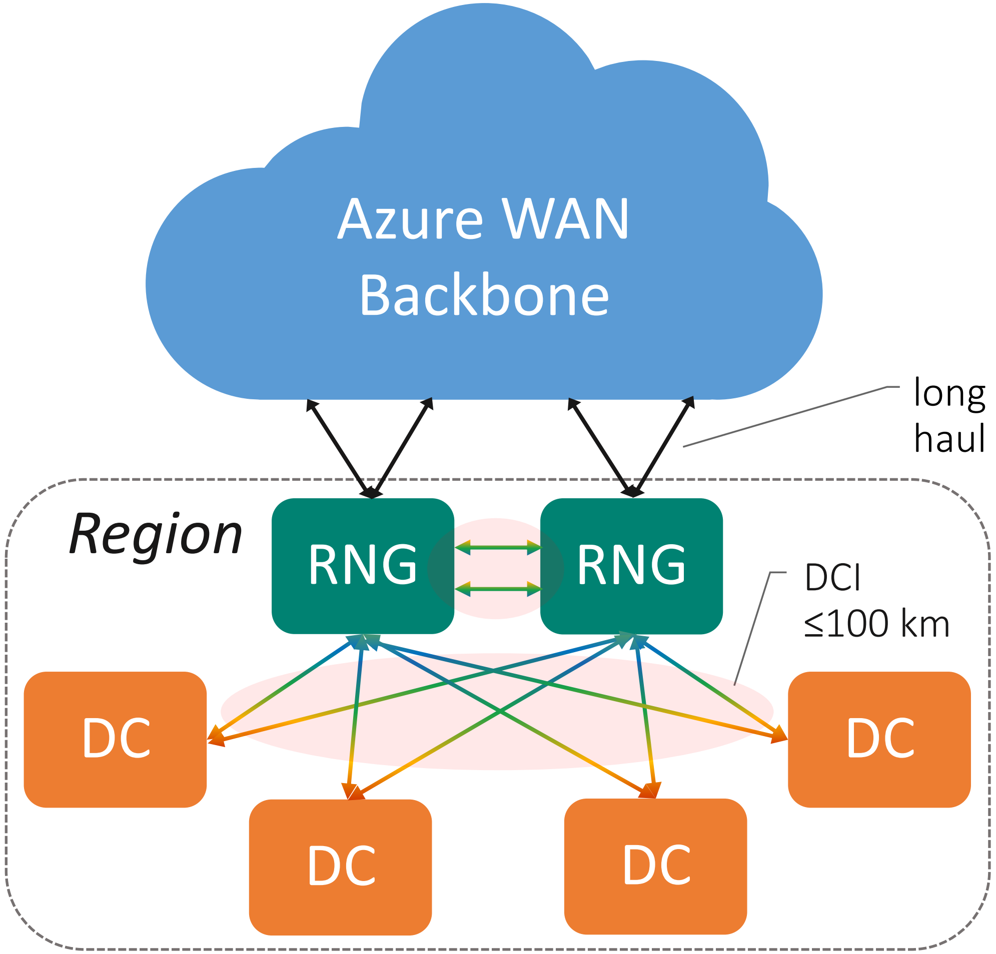
100 Gb/s and 400 Gb/s Coherent Market

- 100 Gb/s and 400 Gb/s market data – coming soon (waiting on approval)

Key Points:

- 400 Gb/s market starts in 2019
- Coherent edge (100G and 400G) represents > 25% of whole coherent market in 2022
    - Edge applications expected to be dominated by Ethernet

Use Case Example – 400Gb/s DCI



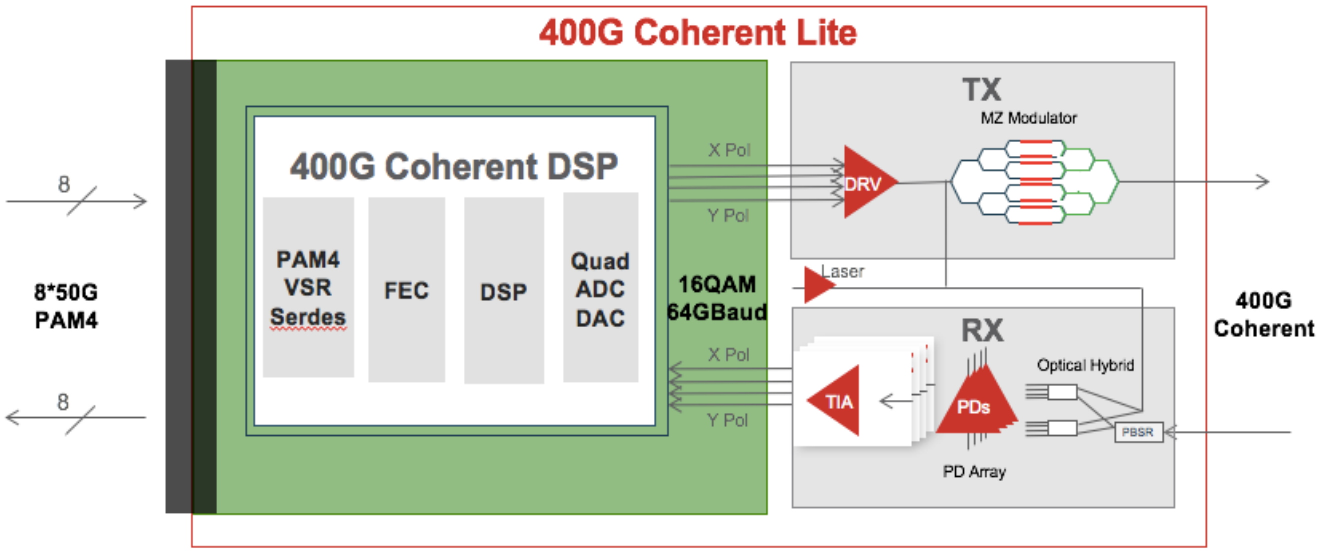
- Distributed data center model
- Massively parallel and highly resilient
- Latency SLAs constrain maximum fiber distances
- Ethernet based fabric
- DWDM based DC interconnect

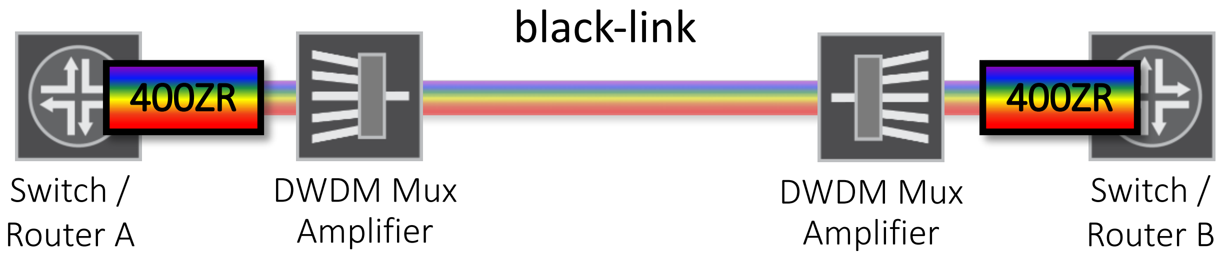
Microsoft

Technical Feasibility

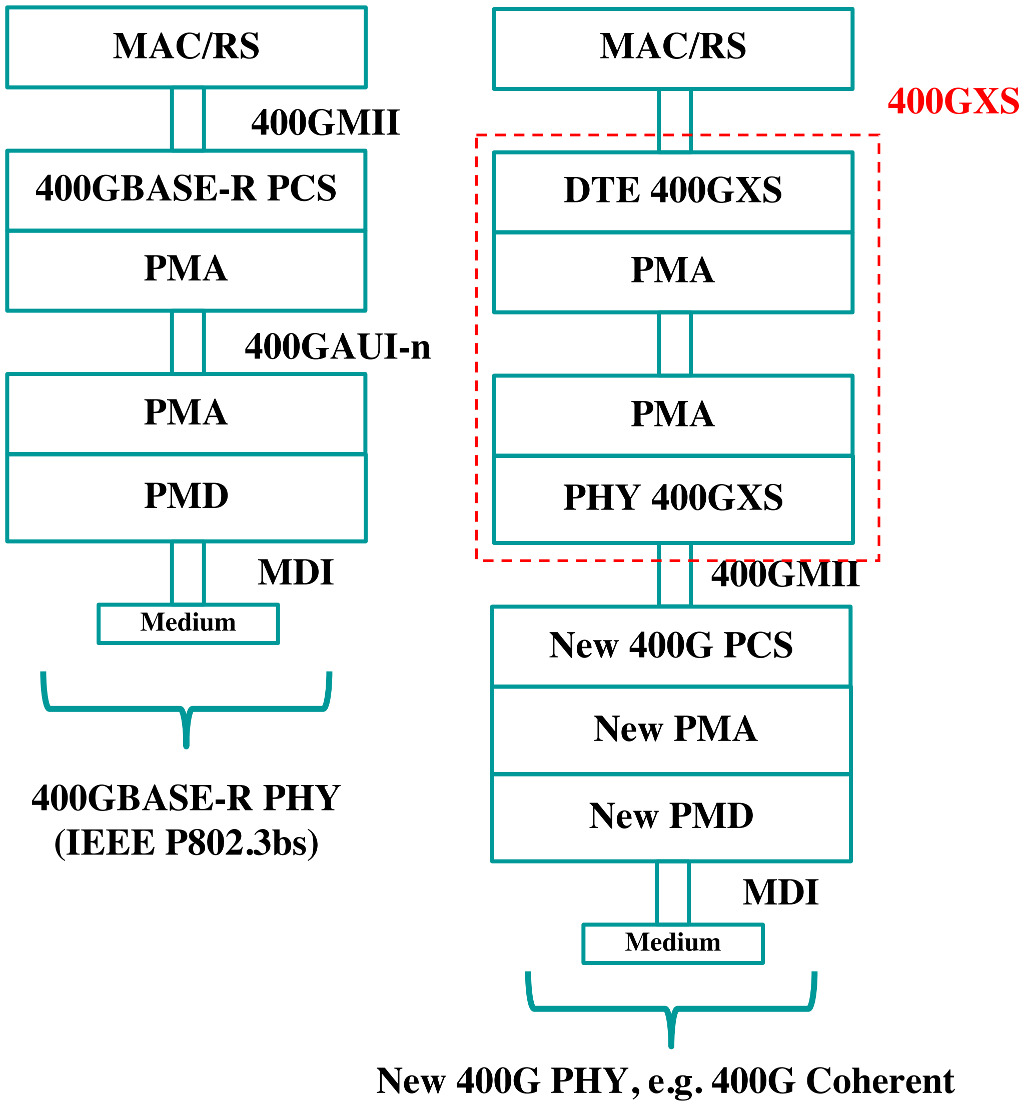
400Gb/s DCI Application Overview

- 400Gb/s Ethernet
- Reach of at least 80km over a DWDM system (black-link)
- Single wavelength
- ≥16 Tb/s per fiber
- Standard 400GAUI electrical interface
- Compatible with 400G Ethernet switch/router ports (i.e. QSFP-DD/OSFP pluggable form factor)
- Multi-vendor interop is critical



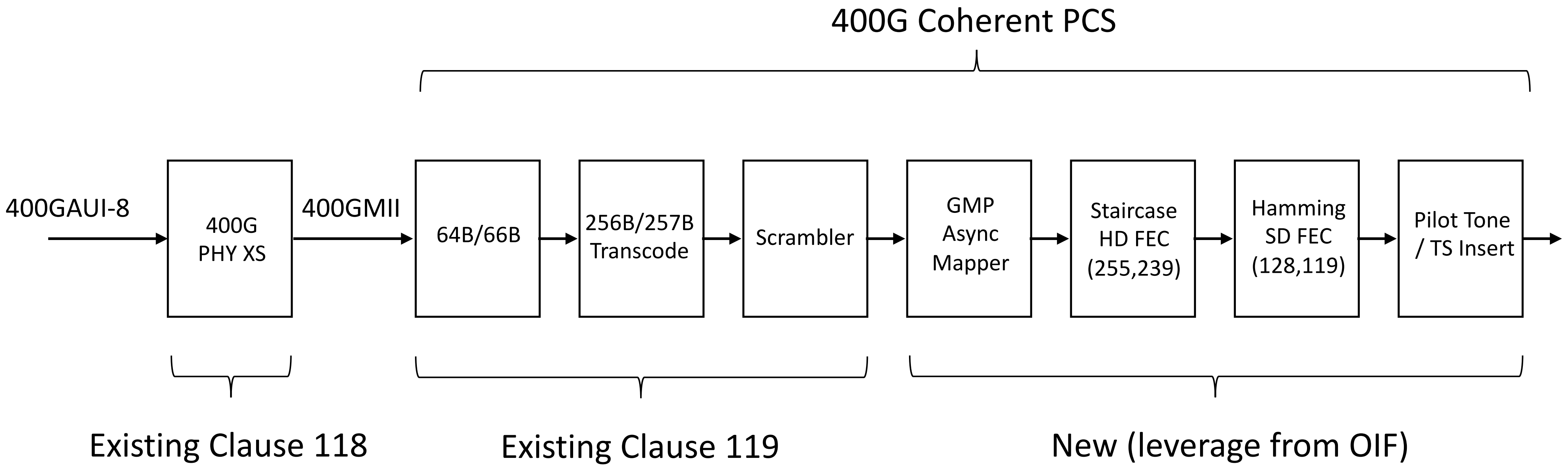


400GbE Extender Sublayer Recap



- 400GXS (Extender Sublayer) extends the 400GMII over a physically instantiated 400GAUI-n
- 400GXS is defined in IEEE P802.3bs Clause 118
- The 400GXS allows a new 400G PHY (with different line coding/FEC) to interface to an existing 400G switch port/ASIC over the 400GAUI
- 400G Coherent DWDM PHY is an example of a new 400G PHY that requires different coding/FEC

400G Coherent PCS Overview



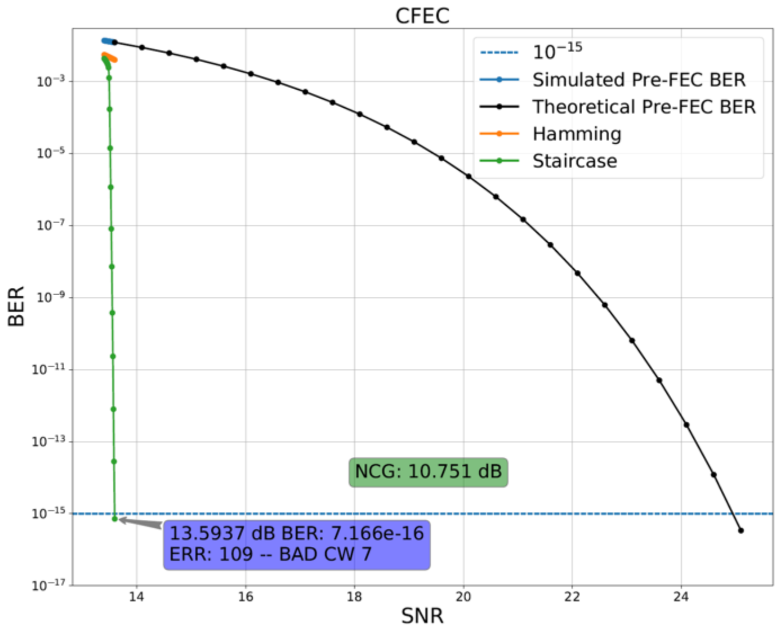
- Reuse significant amount of 802.3bs PCS (Clause 119)
- Leverage FEC from OIF 400ZR project
- Async mapping is something new
    - decouples 400GAUI-8 from the 400G Coherent PMD (different clock domains)
    - ADC & DAC can use same clock reference clock (cost savings in optical module)

400G Coherent FEC

Strong industry consensus already exists for a multi-vendor FEC for 400G coherent applications, and it is currently under development in the OIF as part of the 400ZR project.

400G Concatenated FEC:

- Soft decision inner – Hamming (128,119) Code
- Hard decision outer – Staircase Code (255,239)
- NCG = 10.8dB for 16QAM
- FEC overhead = 14.8 %
- Ultra Low Power = 420 mW (7nm, 400G)
- Burst Tolerance = 1024 bits, including random errors from background noise (more than 2048 bits without background noise)
- Latency = 4 ms



Source Inphi

400G DWDM Coherent Optics

- To be added

Economic Feasibility

- Traditionally long haul transmission technology has been higher power, higher cost and physically larger, than corresponding client interface technology
- With the combination of reduced reach targets (80km) and advanced CMOS technology nodes (7nm), both coherent and client technologies can be implemented in the same form factors
- Economically this significantly lowers network cost and increases flexibility

Comparing 400G Coherent with 400G Client

400G FR4: PAM4 on 4 wavelengths

4 x Laser

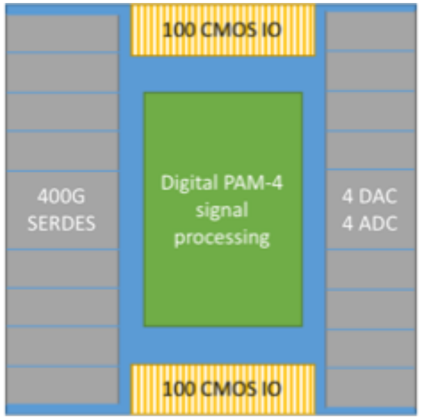
DSP
1 x 400GAUI
4 x ADC
4 x DAC

4 x MZI/EML
4 x Driver

4 x RX
4 x TIA

WDM
Combine
/ splitter

Client DSP



400G Coherent: "PAM4" on XI, XQ, YI, YQ

1 x Laser

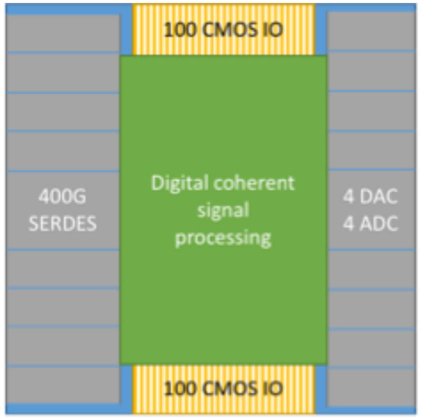
DSP
1 x 400GAUI
4 x ADC
4 x DAC

4 x MZI
4 x Driver

4 x RX
4 x TIA

Pol & IQ
Combine
/ splitter

Coherent DSP



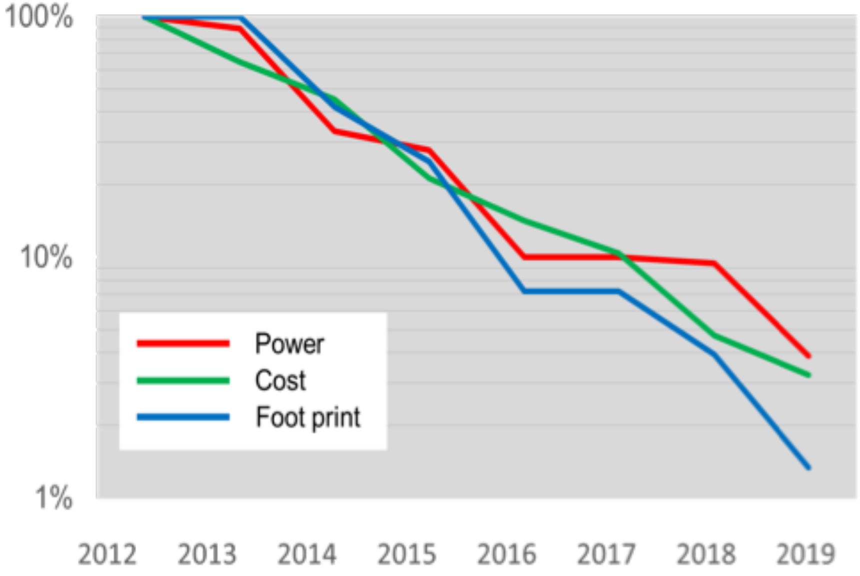
Source Acacia

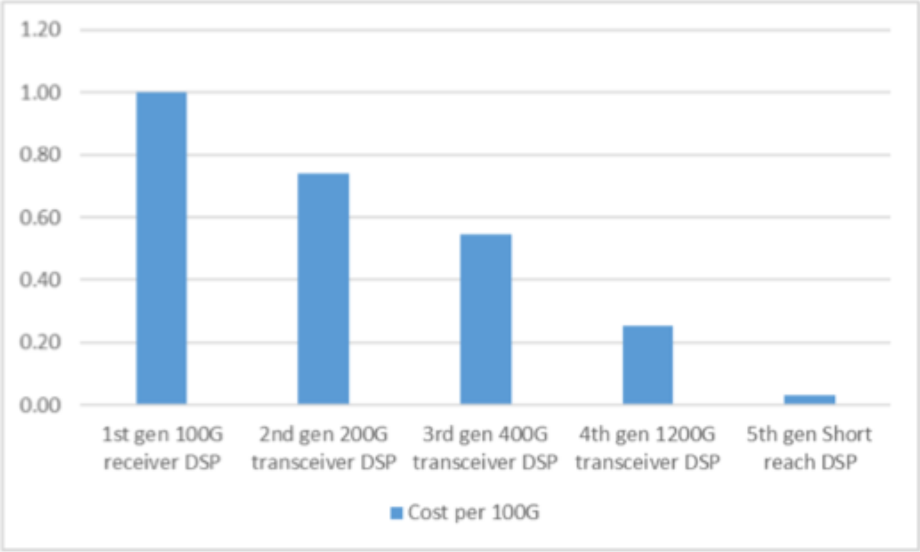
- Similar size die (pad limited)
- > 50% of power consumption is in ADC/DAC and SERDES

Both coherent and client feasible in same form factor

Impact of advanced integration

- Advances in CMOS and optical integration have driven reductions in size, cost and power for coherent solutions





Source Acacia

Summary

- Market need for 400 Gb/s Ethernet DWDM solutions up to 80km has been identified
- Coherent technology for 400 Gb/s long-haul and metro applications exists and is being deployed, suggesting IEEE 802.3 specifications for 80km links are feasible
- Current industry activities and consensus will enable interoperable specifications be developed
- (Hopefully) clear use case identified for proposed SG objective

Proposed Objective

Propose the SG adopts an objective:

- Provide physical layer specifications supporting 400 Gb/s operation on a single wavelength capable of at least 80km over a DWDM system.

Companion objectives

Associated with the proposed PHY objective, the SG would also need the related objectives to be adopted:

- Support a MAC data rate of 400 Gb/s
- Support a BER of better than or equal to 10^-13 at the MAC/PLS service interface (or the frame loss ratio equivalent) for 400 Gb/s

Assumption is these would be all included together in same motion

Backup

DWDM terminology recap

Updated terminology (from 2/27/18 ad hoc)

- **WDM** – optical technology that couples more than one wavelength in the same fiber, thus effectively increasing the aggregate bandwidth per fiber to the sum of the bit rates of each wavelength.
- **DWDM** – A WDM technology where the frequency spacing is less than or equal to 1000 GHz.
- **DWDM PHY**: An Ethernet PHY that operates at a single wavelength on a defined frequency grid and is capable of running over a DWDM system.
- **DWDM Channel**: The transmission path between a DWDM PHY transmitting to another DWDM PHY.
- **DWDM Link**: One DWDM PHY transmitting to one other DWDM PHY through the transmission path between them.
- **DWDM System**: An aggregate of DWDM links over either a single optical fiber or a single optical fiber per direction.
- **DWDM Network** - same as DWDM System so term not to be used
- **In-line amplification:** Optical amplification that resides within a DWDM Channel

Modified/New from ad hoc **Proposed Delete**

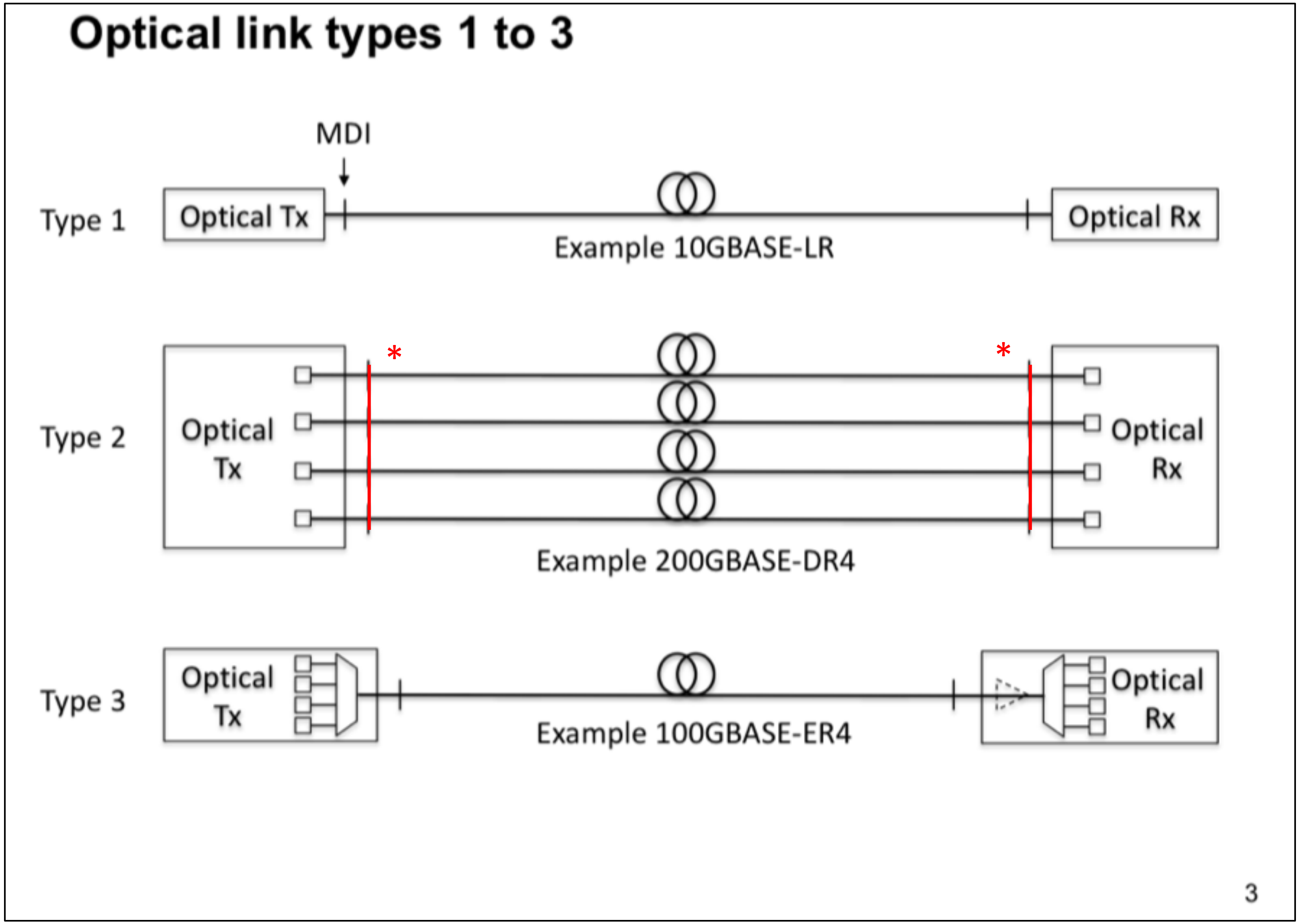
Link Types

Presented by Pete Anslow.

Excellent summary of link type configurations.

Type 1, 2, 3 all represent what would be typical of past IEEE 802.3 PMDs

Common usage would call these “Optical PHYs” as opposed to “Electrical PHYs” and different to the “DWDM PHY” which could be the outcome of the proposed objective in this presentation.



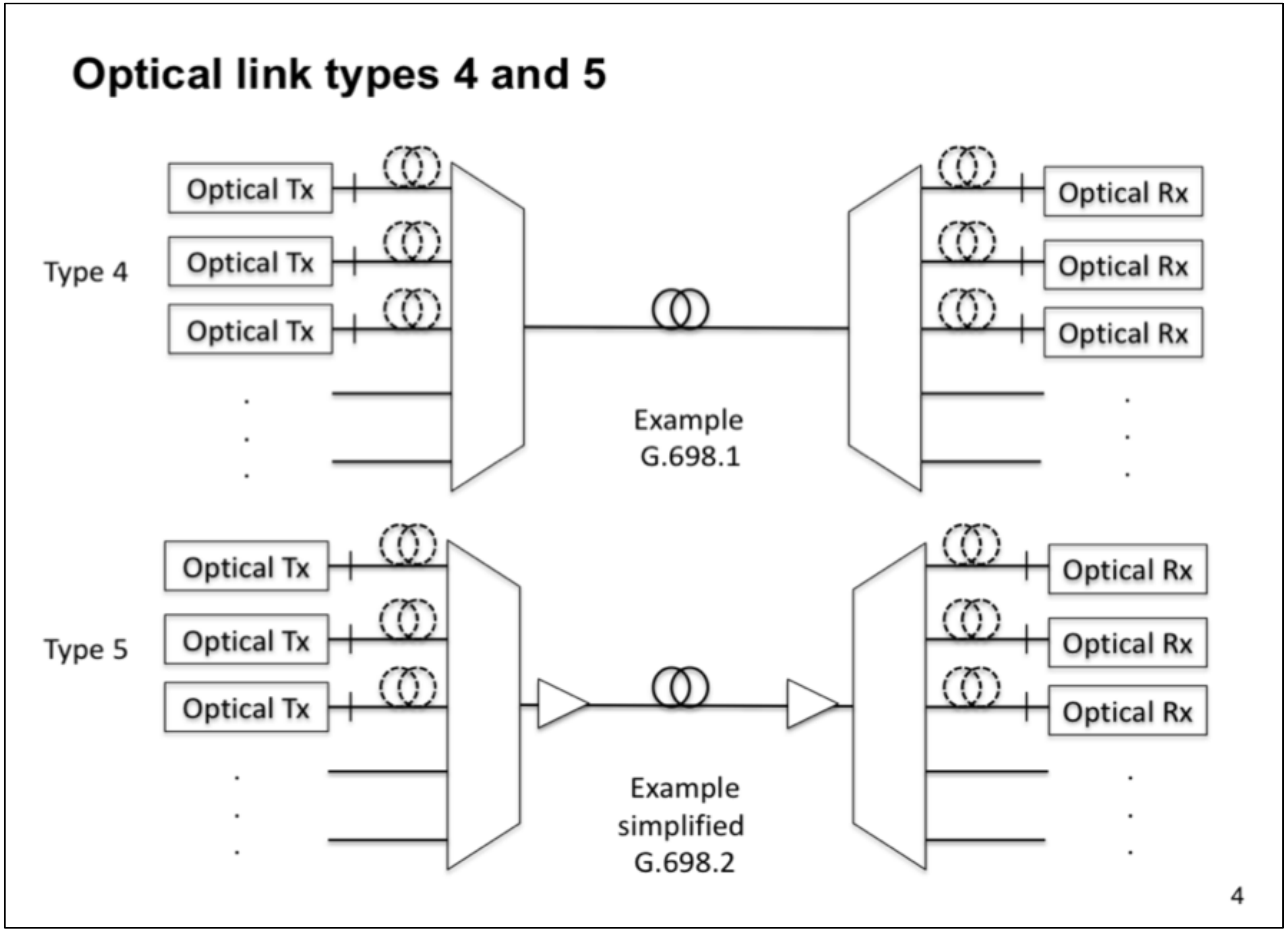
http://www.ieee802.org/3/B10K/public/18_01/anslow_b10k_01_0118.pdf

* Proposed modification to slides

Link Types

Link Types 4 & 5 are representative of network topologies consistent with DWDM Systems and technologies.

The range of Cable/MSO deployments are consistent with both Type 4 & 5 link types



http://www.ieee802.org/3/B10K/public/18_01/anslow_b10k_01_0118.pdf

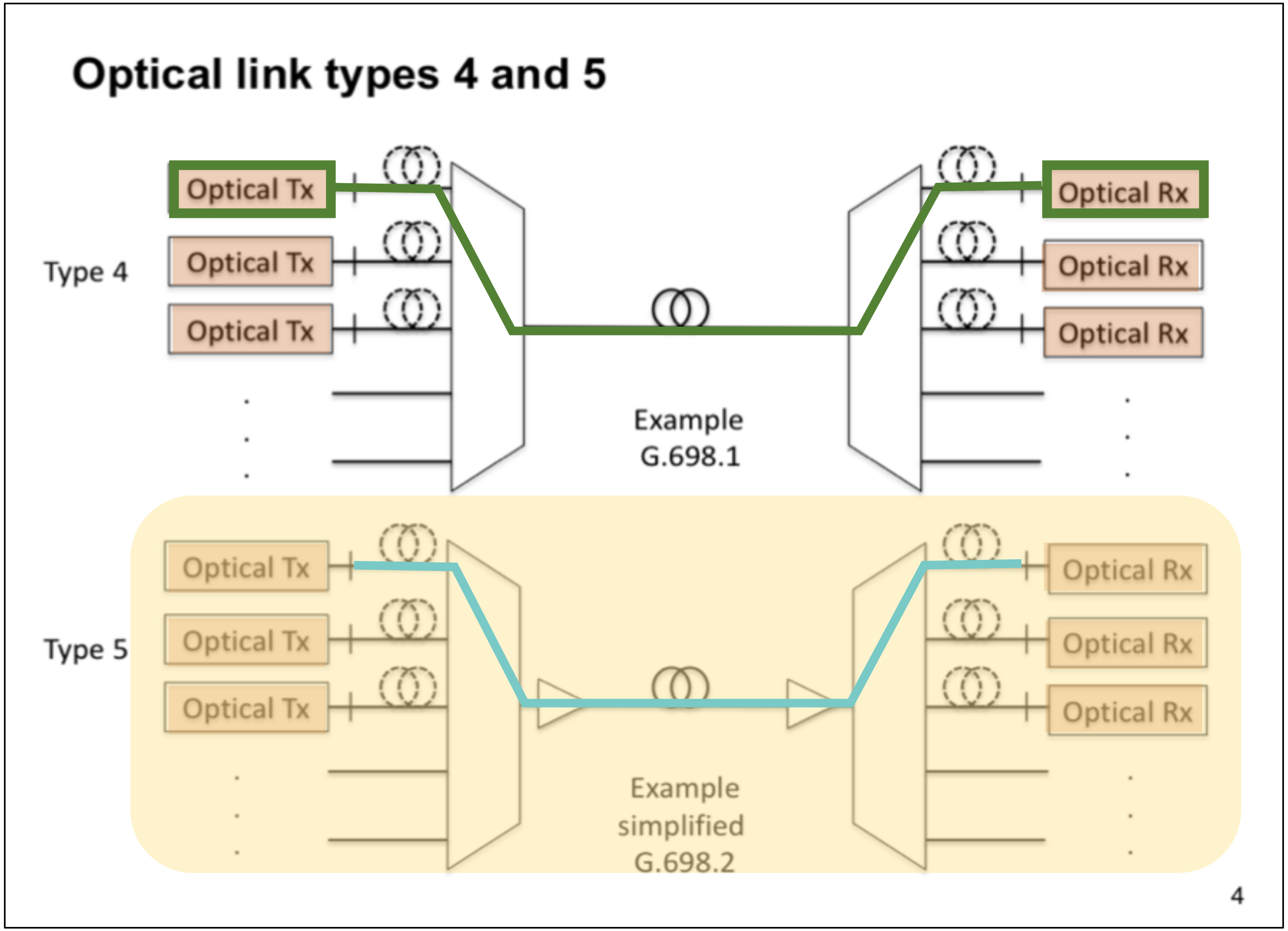
DWDM Link Types and Terminology

DWDM PHY:

DWDM Channel:

DWDM Link:

DWDM System:



http://www.ieee802.org/3/B10K/public/18_01/anslow_b10k_01_0118.pdf

Coherent optical DWDM technology overview

Coherent DWDM Overview

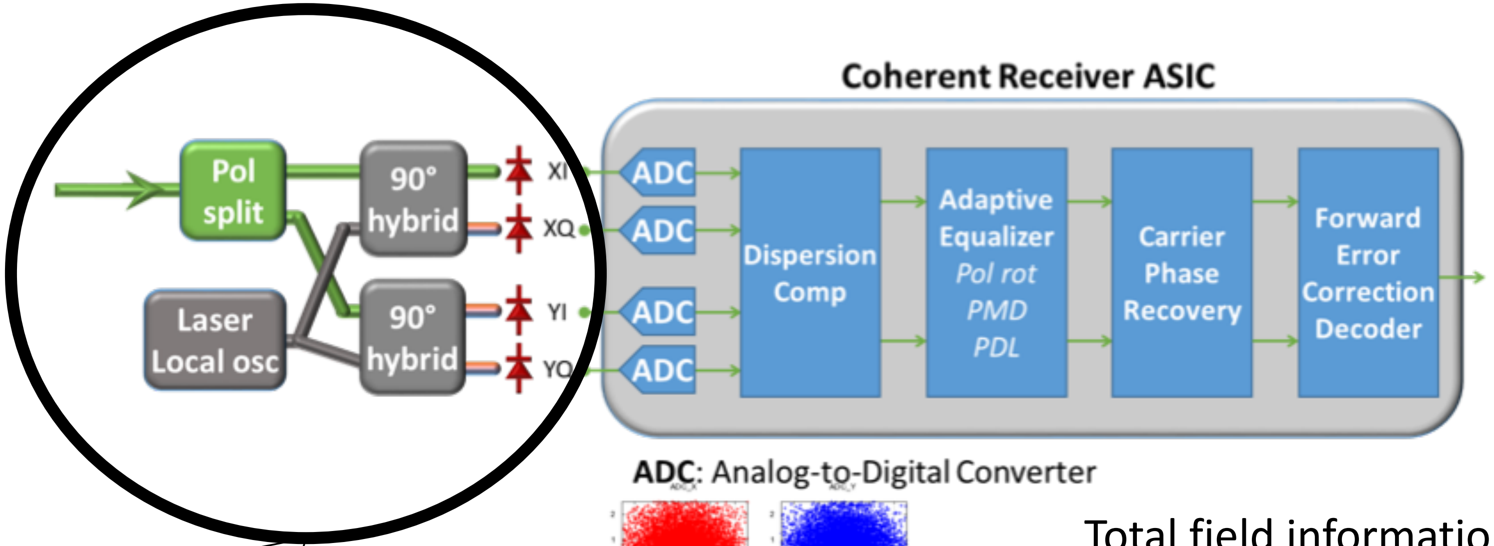
- Coherent optical technology was significantly studied and researched from the mid-80’s due to it’s potential to overcome optical fiber transmission challenges that exist with the direct-detection approaches.
- Invention of Erbium optical amplifier, stalled progress for a while
- Above 10 Gb/s, direct detection transmission was becoming a challenging solution to achieve
- Mid-2000’s, intersection of CMOS and optical technology capabilities opened possibility that a coherent-detection solution was feasible for 40 Gb/s
- March 2008, Nortel (Ciena) announce first commercial transmission system @ 40 Gb/s
- Today, coherent-based transport is now the de-facto standard technology choice for transmission solutions @ 40 Gb/s, 100 Gb/s, 250 Gb/s, 400 Gb/s and beyond
    - Widespread & mature technology
    - Originally targeted for Long-haul and Ultra-long haul solutions, recent market focus includes metro and lower reach optimizations

Key points: Coherent vs. direct detection

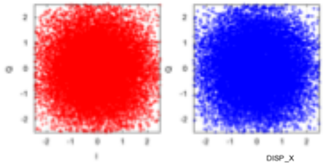
- Coherent transceivers use linear E/O & O/E conversion
  - Use of local oscillator at receiver ensures full optical field (amplitude and phase) survives after the photodetector
  - Enables more complex modulation schemes to be employed to increase capacity
- Linear optical distortions remain linear.
- Digital Signal Processing may then be used to compensate the channel / transceiver
- Complexity / constraints of DSP depends upon application
  - Wide range of optical impairments can be compensated in the DSP
  - Simplifies operational issues
  - Complexity shared between optical and digital technologies

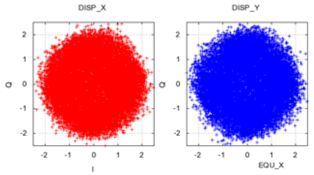
What is Coherent Detection?

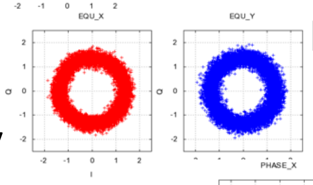
By mixing with a local oscillator, the received optical signal is down-converted to baseband and sampled with a high speed ADC

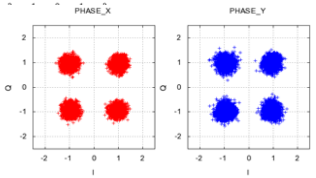


ADC: Analog-to-Digital Converter

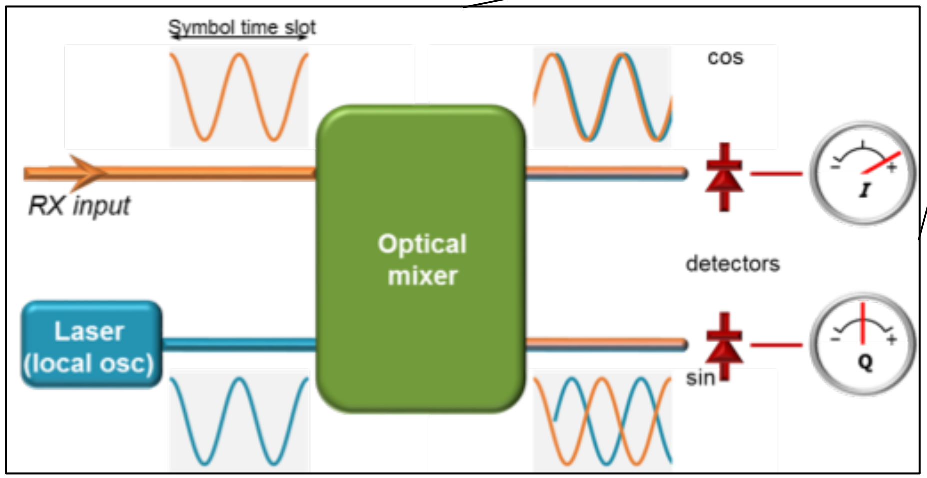








Total field information (amplitude, phase, polarization) is maintained into the electrical/digital domain enabling equalization of link impairments by a digital signal processor (DSP)

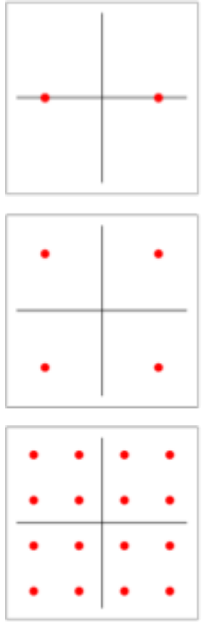


Coherent detection requires the addition of a local oscillator (usually shared with transmit laser) and additional DSP processing. Effectively amplifies the incoming signal

How Does Coherent Increase Capacity?

Modulation in Phase and Amplitude

Capacity



BPSK
2 (=2^1) states: 1 bit pr symbol ("time slot")
Modulating Signal – 1 NRZ

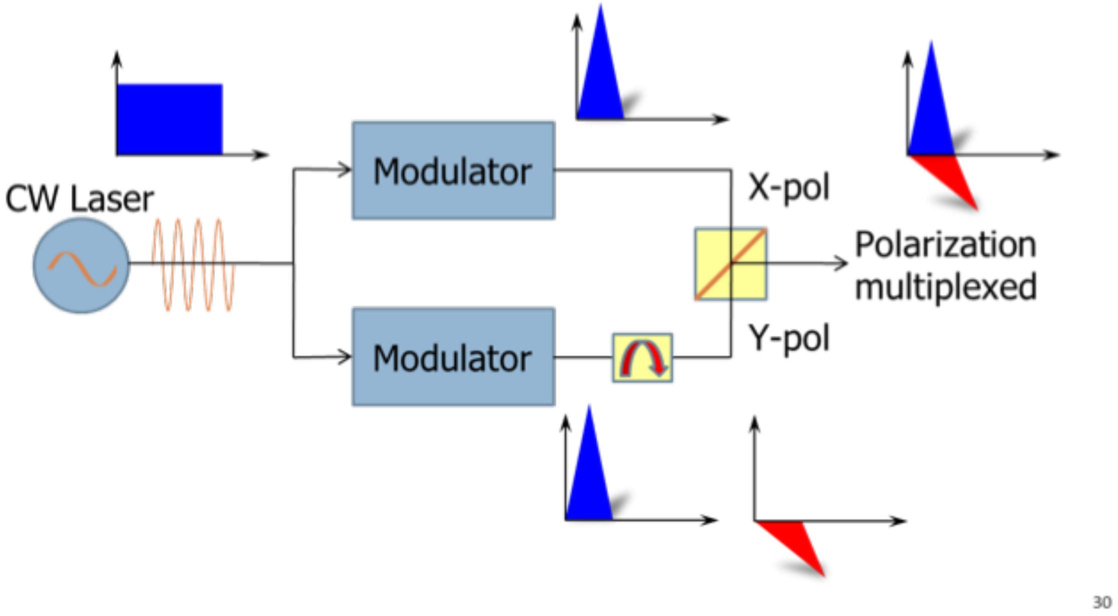
QPSK
4 (=2^2) states: 2 bit pr symbol ("time slot")
Modulating Signal – 2 NRZ (I/Q)

16QAM
16 (=2^4) states: 4 bit pr symbol ("time slot")
Modulating Signal – 2 PAM4 (I/Q)

Performance

2x capacity of intensity modulation

Polarization Multiplexing



2x capacity of single polarization

Coherent DWDM Applications

Proven Applications

Aspect	Cable/MSO/ Data Center	Metro	Regional	Long-Haul	Submarine
Reach (km)	80	300	600	4000	10000
Chromatic Dispersion (ps/nm)	1280	5000	15000	80000	240000
DGD max (ps)	16	30	43	111	35
Latency (Critical?)	Yes	Sometimes	Less so	Not so	Not really
OSNR/FEC	Low Perf.	Hi-perf	Hi-perf	Hi-perf	Hi-perf
Cost	Low	Low	Mid	Mid	High
Power	Ultra low	Low	Mid	Mid	Mid