IP over DWDM

(What IP folks should know)

Tashi Phuntsho | Technology Evangelist @FLEXOPTIX

IP-over-Transport: The Evolution

T1/E1 circuits

Early IP over synchronous serial lines.

Typical framing:
PPP in HDLC-like framing over T1/E1

SDH/SONET

IP packets over standardized transport.

Common framing:
POS (PPP over SONET/SDH)

DWDM + OTN

Capacity scaled with wavelengths (mux/transponders)

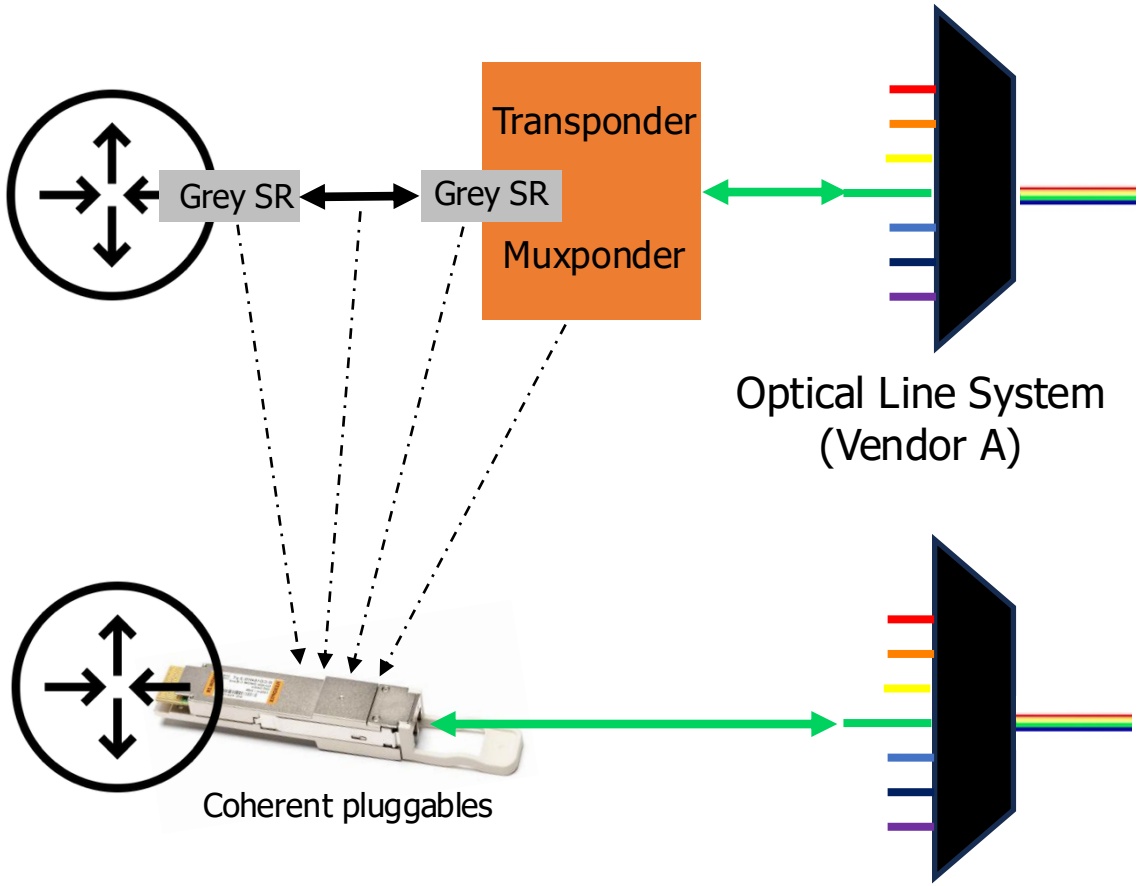
Common mapping:
Ethernet/packet into OTN, over DWDM

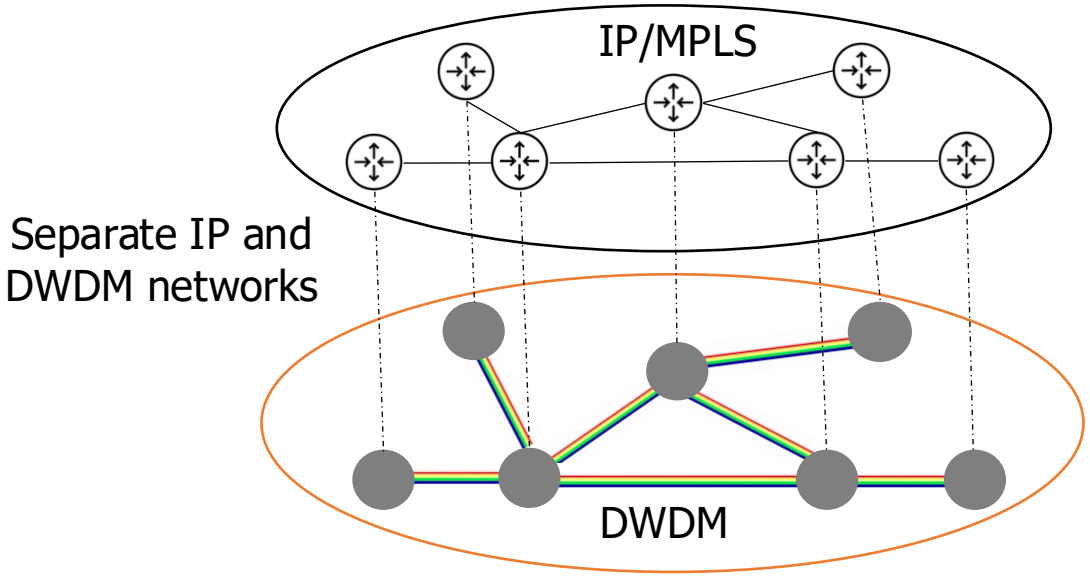
IPoDWDM

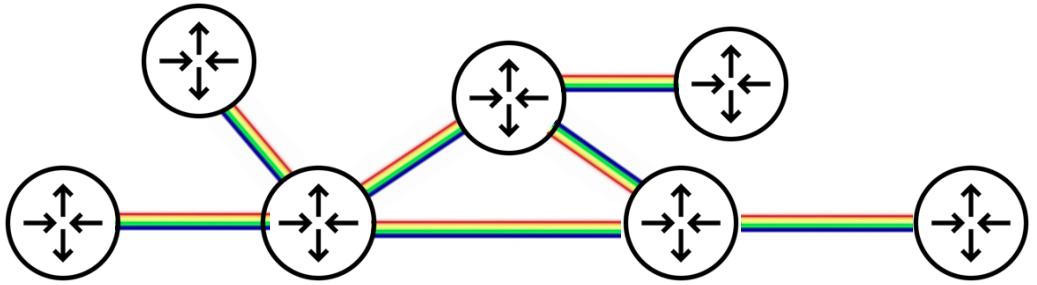
Coherent optics directly in routers and switches.

Model:
Ethernet/IP directly on wavelengths

IP-over-DWDM: The Concept







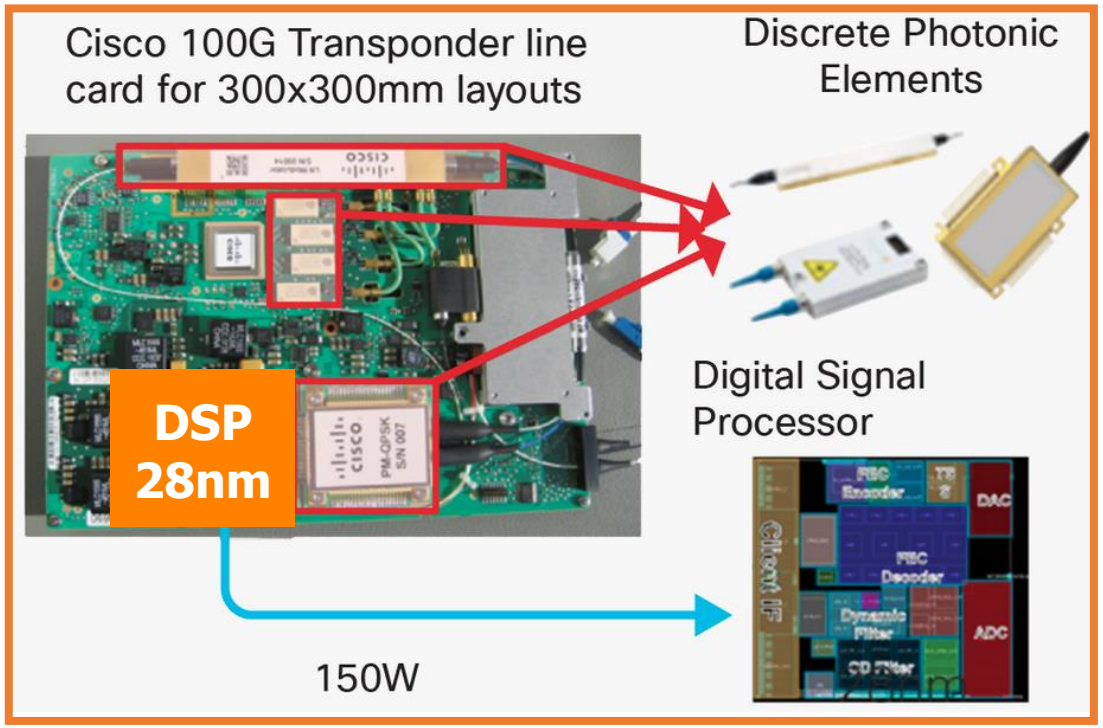
Simplified single packet/routed optical network

Recent Operator Adoptions

Network	Type	Speed	Span	Status	Source
Internet2	R&E	400G backbone links	600 Km	Production	Internet2
Arelion	CSP	400G ZR+	675 Km	Production	Arelion
RETN	CSP	400G backbone links	300Km ~ 950 Km	Production	RETN
Colt	CSP	800G ZR+	667 Km	Trial	Colt
GÉANT	R&E	400G ZR++ ULH	3400 Km	Trial	GÉANT
AARNet	R&E	400G ZR++ ULH	4600 Km	Trial	AARNet
BKNIX	IXP	400G ZR+	40Km	Unamplified DCI	
DrukREN	R&E	400G ZR+	96Km	Amplified PtP	
Network-A	CSP	400G ZR+	196Km & 105Km	Protection link – WIP (Nepal)	
Network-B	CSP	100G ZR+ 400GZR+	190Km & 60Km (ROADM) 115Km (ROADM)	WIP (Mongolia)	

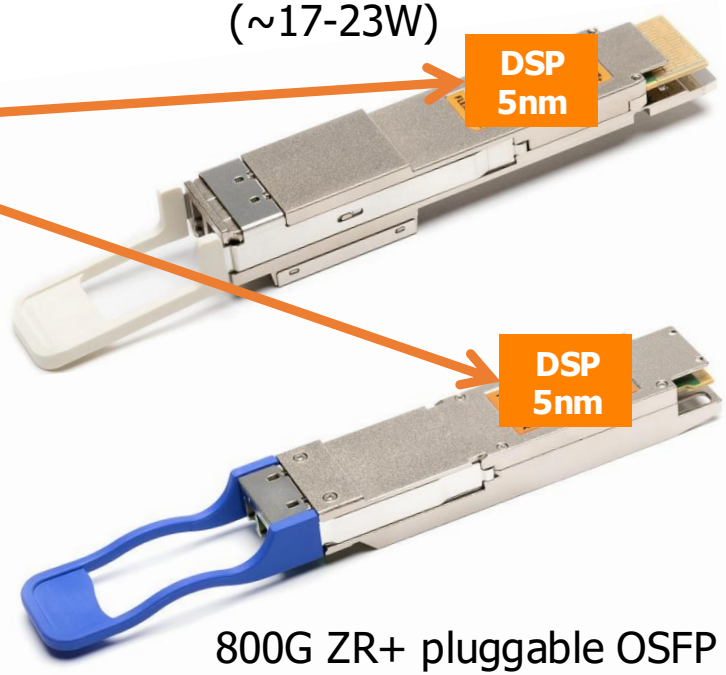
What makes it possible?

Coherent optics became **Pluggable**



Source: Cisco Systems

400G ZR/ZR+ pluggable QSFP-DD (~17-23W)



800G ZR+ pluggable OSFP (~29W)

- Router/switch friendly xSFP form factors
- Reduced power-consumed per-bit

Open Standards & Disaggregation

OIF 400ZR

- Standardised coherent interfaces across vendors ~ interoperable ecosystem
- Good for short reach (80Km), single-span links for Data Center Interconnect (DCI)

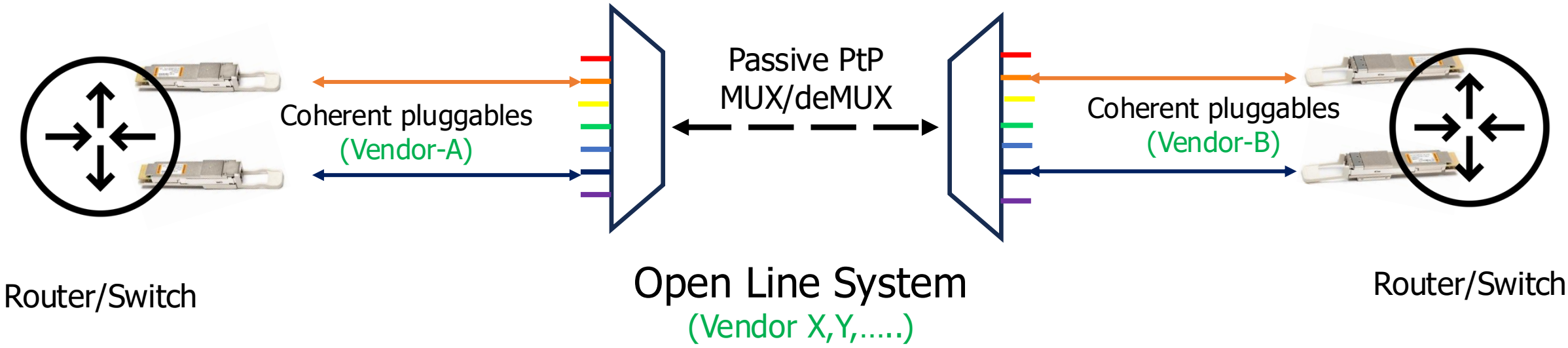
OpenROADM

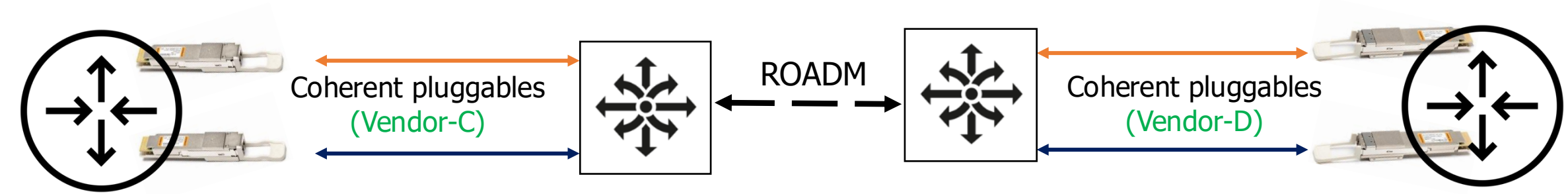
- Push for open line systems and remove restrictions on alien wave support
- Reduce dependence on closed single vendor optical stacks ~ disaggregated transport

OpenZR+ MULTI-SOURCE AGREEMENT

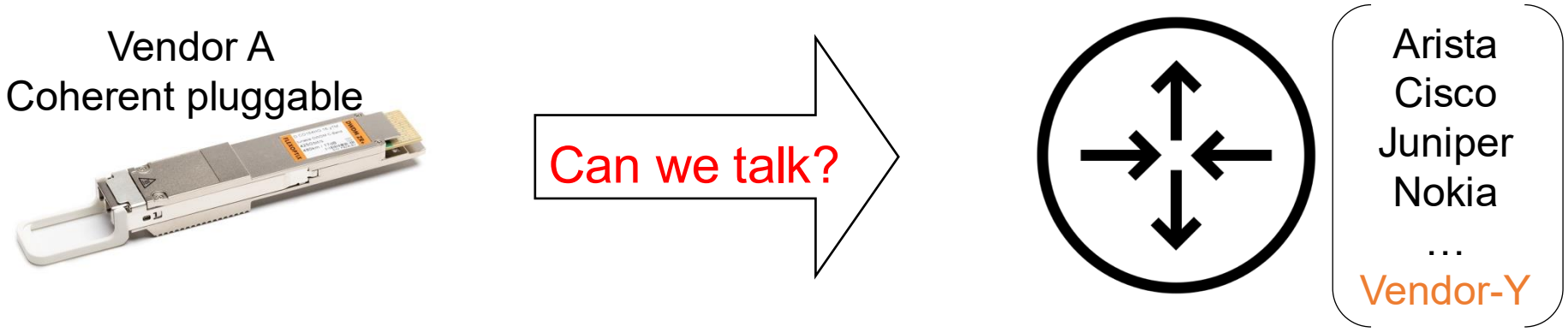
- Extended reach (beyond 120Km) with high-performance FEC (oFEC)
- Flexibility: multi-rate options (trade capacity for reach)

Disaggregated & Open IPoDWDM





Interoperability with Hosts



C-CMIS: Coherent Common Management Interface Specification

- Defined by OIF (CMIS v5.3 & C-CMIS v1.4)
- Defines how coherent plugs communicate with hosts for management and control
- Helps make coherent plugs more portable across different host platforms

Is it just plug & play?

Plug the coherent optics into a router/switch:

- (If) Form factor fits
- (If) Host can control/manage the plug
- Why wouldn't it just work automagically? hmm…

- Does a simplified IPoDWDM architecture remove optical engineering challenges?

Let us look at the optical telemetry exposed by Routers and Switches:

- What do the stats convey?
- How do they affect the link?
  - Distance/reach?
  - Capacity?

Parameter	MIN	AVG	MAX
CD(Short)[ps/nm]	1093.00	1093.00	1093.00
DGD[ps]	1.00	1.00	1.00
RX PWR[dBm]	-9.22	-9.14	-8.97
TX PWR[dBm]	-1.04	-1.00	-0.97
OSNR[dB]	23.10	23.38	23.40
RX CHAN PWR[dBm]	-13.26	-13.05	-12.82
ESNR[dB]	14.30	14.41	14.50
LASER BIAS[mA]	278.60	279.10	279.29
FREQ OFF[Mhz]	162.00	257.48	346.00
SOP RATE[krad/s]	0.00	0.00	0.00
PDL[dB]	0.60	0.76	1.00
SOPMD[ps^2]	26.62	47.59	72.06

EC BITS : 0
UC WORDS : 0

Parameter	MIN	AVG	MAX
PREFEC BER	8.50e-03	8.75e-03	9.60e-03
POSTFEC BER	0.00e+00	0.00e+00	0.00e+00
Q FACTOR[dB]	7.30	7.46	7.50
Q MARGIN[dB]	0.80	0.90	1.00

Example:
400G ZR+ on Cisco N93600CD-GX (DrukREN)

Parameter	MIN	AVG	MAX
CD(Short)[ps/nm]	1093.00	1093.00	1093.00
DGD[ps]	1.00	1.00	1.00
RX PWR[dBm]	-9.22	-9.14	-8.97
TX PWR[dBm]	-1.04	-1.00	-0.97
OSNR[dB]	23.10	23.38	23.40
RX CHAN PWR[dBm]	-13.26	-13.05	-12.82
ESNR[dB]	14.30	14.41	14.50
LASER BIAS[mA]	278.60	279.10	279.29
FREQ OFF[Mhz]	162.00	257.48	346.00
SOP RATE[krad/s]	0.00	0.00	0.00
PDL[dB]	0.60	0.76	1.00
SOPMD[ps^2]	26.62	47.59	72.06

EC BITS : 0
UC WORDS : 0

Parameter	MIN	AVG	MAX
PREFEC BER	8.50e-03	8.75e-03	9.60e-03
POSTFEC BER	0.00e+00	0.00e+00	0.00e+00
Q FACTOR[dB]	7.30	7.46	7.50
Q MARGIN[dB]	0.80	0.90	1.00

CD/DGD: fiber impairments that causes errors - the DSP needs to compensate for.

RX PWR: is the received light weak, good, or too bright?

TX PWR: did the light launch correctly?

OSNR/ESNR: How clean is the signal? Higher is better!

Pre-FEC BER: does the DSP need to correct errors?

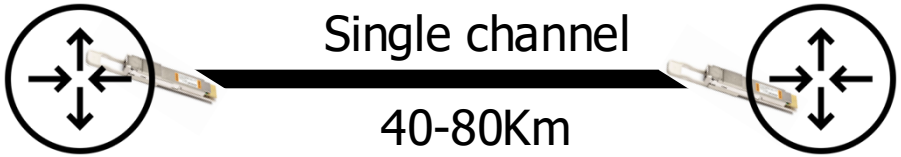
Post-FEC BER: are errors leaking through the correction?

Q-factor/Q-margin: how far/hard can we push before link fails?

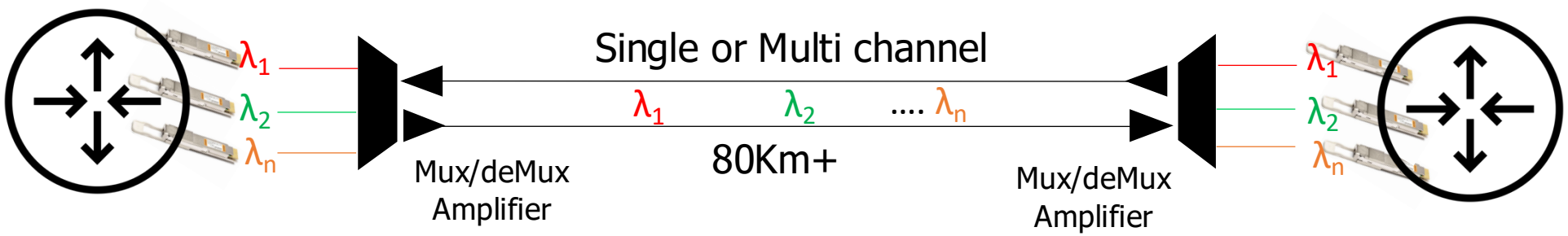
Laser Bias: is the laser aging? Is there a temperature stress?

Frequency Offset: is the wavelength sitting where it is expected on the grid?

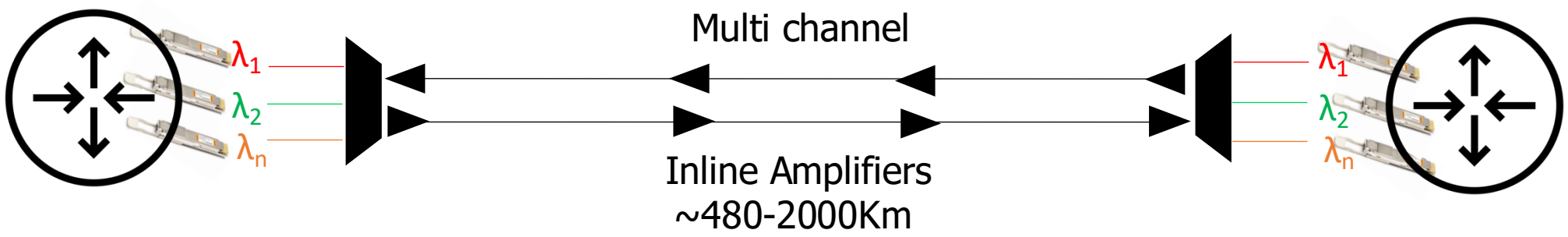
Pluggable for all requirements?



Unamplified
Dark Fiber



Amplified
point-to-point
DWDM



Full featured
DWDM systems

Use Case1: Short single span DCI

Link:

- Single channel
- Unamplified dark fiber
- 40Km link

What was measured?

```
transceiver {
        functional-type digital-coherent-optics
        ……
        ethernet-pmd 400G-ZR-Amp
        connector-type LC
        vendor FLEXOPTIX
        vendor-part-number D.CO164HG.16.yTN
        ……..
        supported-operational-mode [
            1
            2
            3
            5
            6
            10
            11
        ]
```

```
optical-channel 1 {
            frequency 193100000
            operational-mode 10
            target-power 1.00
            laser-tunability flex-tunable
            oper-frequency 193700000
            minimum-frequency 191300000
            maximum-frequency 196100000
            rx-los-reaction squelch
            rx-los-thresh -23.00
            module-state ready
            rx-electrical-snr-x-polarization 18.4
            rx-electrical-snr-y-polarization 18.4
            rx-quality-margin 4.8
            rx-optical-snr-x-polarization 29.2
            rx-optical-snr-y-polarization 29.2
```

```
channel 1 {
            wavelength 1552.52
            input-power {
                latest-value -9.59
                …
            }
            output-power {
                latest-value 4.01
                ….
            }
```

Module range for the selected app-code:

- RX= -22dBm to 0dBm
- TX = +3dBm to +5dBm
- OSNR = 34dB

Hmm… Operational Mode?

Coherent transceivers support multiple config options

- Baud rate, modulation scheme, FEC, channel spacing, CD thresholds, etc
- aka ‘application codes’

Application Code	Host Interface	Baud Rate [GBd]	Modulation	Bandwidth	FEC	Media Lane Reach	Frequency Range DWDM Grid	Grid Spacing	Flexible DWDM Grid	Tx Power (adjustable)	Rx Overload	Rx Input Range	Coherent Standard	Rx OSNR Tolerance	Rx Sensitivity @ OSNR	CD Tolerance max @0.5 dB OSNR Penalty	DGD Max (PMD tolerance)	Pre-FEC BER Limit	Post-FEC BER
AppCode 1	400GAUI-8	59.843	DP-16QAM	400GbE	cFEC	120 km (amplified)	191.3 - 196.1 THz (C-band)	100 / 75 / 50 GHz	6.25 GHz	-10 dBm to +1 dBm	+10 dBm	-12 dBm to 0 dBm	400G ZR DWDM	>26 dB/0.1 nm	-12 dBm OSNR >26dB	±2,400 ps/nm	33 ps	~1.22E-2	<1E-15
AppCode 2	400GAUI-8	59.843	DP-16QAM	400GbE	cFEC	80km (unamplified)	193.7 THz (fixed)	N/A (single λ)	N/A (single λ)	+3 dBm to +5 dBm	+10 dBm	-21 dBm to 0 dBm	400G ZR Grey, unamplified	>26 dB/0.1 nm	-20 dBm (400ZR) OSNR >34dB	±2,400 ps/nm	33 ps	~1.22E-2	<1E-15
AppCode 3	4x100GAUI-2	59.843	DP-16QAM	4x100GbE	cFEC	120km (amplified)	191.3 - 196.1 THz (C-band)	100 / 75 / 50 GHz	6.25 GHz	-10 dBm to +1 dBm	+10 dBm	-12 dBm to 0 dBm	4*100G ZR DWDM	>26 dB/0.1 nm	-12 dBm OSNR >26dB	±2,400 ps/nm	33 ps	~1.22E-2	<1E-15
AppCode 4	4x100GAUI-2	59.843	DP-16QAM	4x100GbE	cFEC	80km (unamplified)	193.7 THz (fixed)	N/A (single λ)	N/A (single λ)	+3 dBm to +5 dBm	+10 dBm	-21 dBm to 0 dBm	4*100G ZR Grey, unamplified	>26 dB/0.1 nm	-20 dBm (400ZR) OSNR >34dB	±2,400 ps/nm	33 ps	~1.22E-2	<1E-15
AppCode 5	400GAUI-8	60.138	DP-16QAM	400GbE	oFEC	480km (amplified)	191.3 - 196.1 THz (C-band)	100 / 75 / 50 GHz	6.25 GHz	-10 dBm to +1 dBm	+10 dBm	-12 dBm to 0 dBm	400G ZR+ DWDM	>24 dB/0.1 nm	-12 dBm OSNR >24dB	±20,000 ps/nm	66 ps	1.7E-2 – 2.0E-2	<1E-15
AppCode 6	4x100GAUI-2	60.138	DP-16QAM	4x100GbE	oFEC	480km (amplified)	191.3 - 196.1 THz (C-band)	100 / 75 / 50 GHz	6.25 GHz	-10 dBm to +1 dBm	+10 dBm	-12 dBm to 0 dBm	4*100G ZR+ DWDM	>24 dB/0.1 nm	-12 dBm OSNR >24dB	±20,000 ps/nm	66 ps	1.7E-2 – 2.0E-2	<1E-15
AppCode 7	3x100GAUI-2	60.138	DP-8QAM	3x100GbE	oFEC	600km (amplified)	191.3 - 196.1 THz (C-band)	100 / 75 / 50 GHz	6.25 GHz	-10 dBm to +1 dBm	+10 dBm	-15 dBm to 0 dBm	3*100G ZR+ DWDM	>21 dB/0.1 nm	-15 dBm OSNR >21dB	±40,000 ps/nm	83 ps	1.7E-2 – 2.0E-2	<1E-15
AppCode 8	2x100GAUI-2	60.138	DP-QPSK	2x100GbE	oFEC	1000km (amplified)	191.3 - 196.1 THz (C-band)	100 / 75 / 50 GHz	6.25 GHz	-10 dBm to +1 dBm	+10 dBm	-18 dBm to 0 dBm	2*100G ZR+ DWDM	>16 dB/0.1 nm	-18 dBm OSNR >16dB	±50,000 ps/nm	83 ps	1.7E-2 – 2.0E-2	<1E-15
AppCode 9	1x100GAUI-2	30.069	DP-QPSK	1x100GbE	oFEC	2000km (amplified)	191.3 - 196.1 THz (C-band)	100 / 75 / 50 GHz	6.25 GHz	-10 dBm to +1 dBm	+10 dBm	-18 dBm to 0 dBm	1*100G ZR+ DWDM	>12.5 dB/0.1 nm	-18 dBm OSNR >12.5dB	±100,000 ps/nm	100 ps	1.7E-2 – 2.0E-2	<1E-15
AppCode 10	400GAUI-8	60.138	DP-16QAM	400GbE	cFEC	90km (unamplified)	193.7 THz (fixed)	N/A (single λ)	N/A (single λ)	+3 dBm to +5 dBm	+10 dBm	-22 dBm to 0 dBm	400G ZR+ grey, unamplified	>34 dB/0.1 nm	-22 dBm OSNR >34dB	±20,000 ps/nm	66 ps	1.7E-2 – 2.0E-2	<1E-15
AppCode 11	4x100GAUI-2	60.138	DP-16QAM	4x100GbE	cFEC	90km (unamplified)	193.7 THz (fixed)	N/A (single λ)	N/A (single λ)	+3 dBm to +5 dBm	+10 dBm	-22 dBm to 0 dBm	4*100G ZR+ grey, unamplified	>34 dB/0.1 nm	-22 dBm OSNR >34dB	±20,000 ps/nm	66 ps	1.7E-2 – 2.0E-2	<1E-15

Hmm.... Operational Mode?

The idea is to select the relevant app code or op mode based on network requirements:

- Avoids having to configure parameters through the CLI
- Ensure the module works correctly: host side vs optics side mapping

JunOS

```
set interfaces <int> optics-options application [options]
```

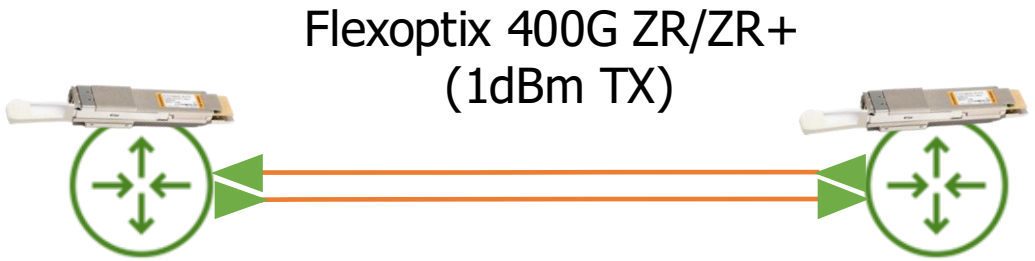
IOS XR

```
(config)#controller optics <interface>
((config-Optics))#appsel simple code <app-code>
((config-Optics))#commit
```

Use Case2: Amplified Point-to-Point links

Setup:

- Link: ~ 100Km OPGW
- Booster & pre-amp: Packet Light PL-1000IL
- Router: Cisco N93600CD-GX
- Single Channel (no Mux/deMux)



What was measured?

CD: 1156 ps/nm ~ 96.9Km link (~20000ps/nm)

RX Power: -9.3dBm (threshold -12dBm)

OSNR: 25dB (threshold 24dB)

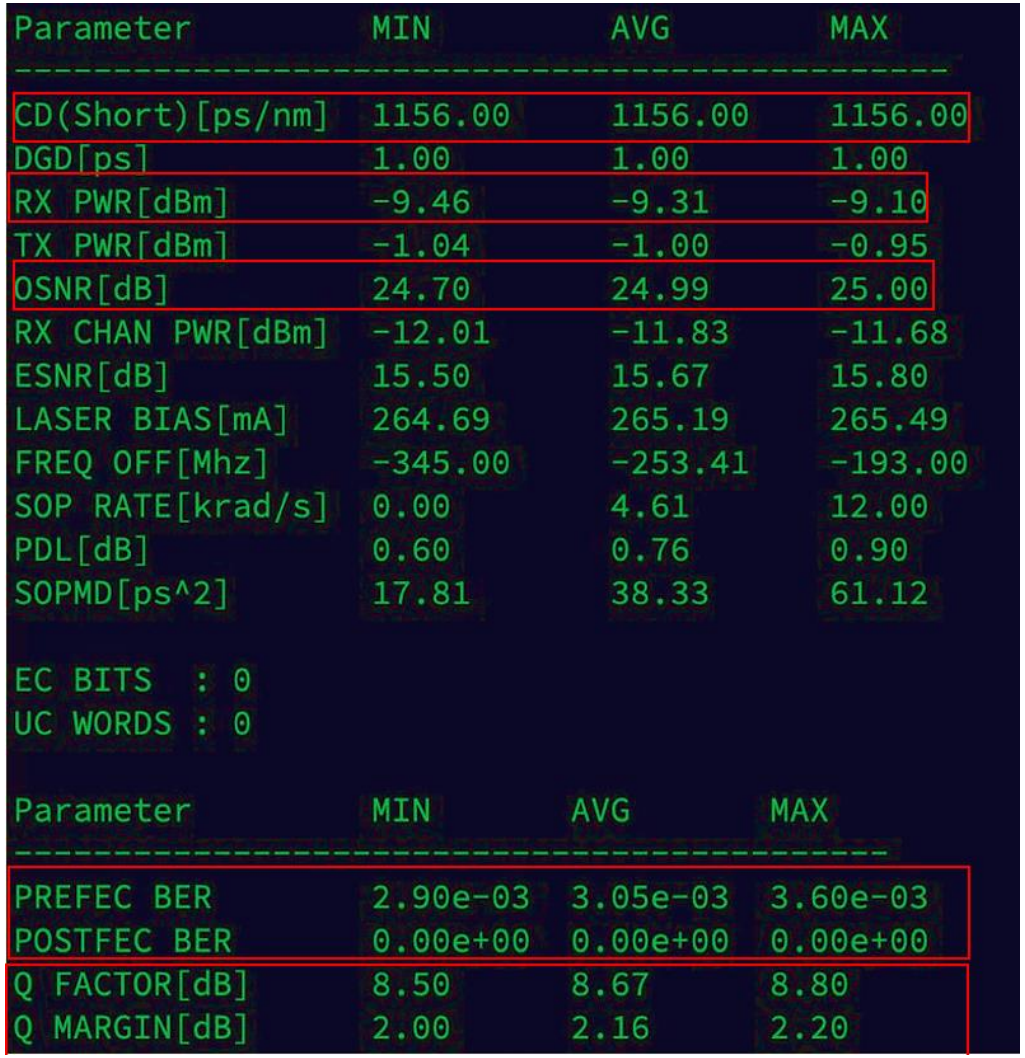
- Enough signal quality (1dB close for comfort)

Pre-FEC BER: Errors received but correctable

Q-margin: 2dB (small margin before link fails)

Review:

- Bad splices on the link?
- Connector losses higher than usual?
- If path loss was ~ 25dB (current ~28dB):
  - Pre-amp input: approx -16dBm (currently -19dBm)
  - OSNR: could improve to 27-28dB



Parameter	MIN	AVG	MAX
CD(Short)[ps/nm]	1156.00	1156.00	1156.00
DGD[ps]	1.00	1.00	1.00
RX PWR[dBm]	-9.46	-9.31	-9.10
TX PWR[dBm]	-1.04	-1.00	-0.95
OSNR[dB]	24.70	24.99	25.00
RX CHAN PWR[dBm]	-12.01	-11.83	-11.68
ESNR[dB]	15.50	15.67	15.80
LASER BIAS[mA]	264.69	265.19	265.49
FREQ OFF[Mhz]	-345.00	-253.41	-193.00
SOP RATE[krad/s]	0.00	4.61	12.00
PDL[dB]	0.60	0.76	0.90
SOPMD[ps^2]	17.81	38.33	61.12

EC BITS : 0
UC WORDS : 0

Parameter	MIN	AVG	MAX
PREFEC BER	2.90e-03	3.05e-03	3.60e-03
POSTFEC BER	0.00e+00	0.00e+00	0.00e+00
Q FACTOR[dB]	8.50	8.67	8.80
Q MARGIN[dB]	2.00	2.16	2.20

Use Case2: Amplified Point-to-Point links

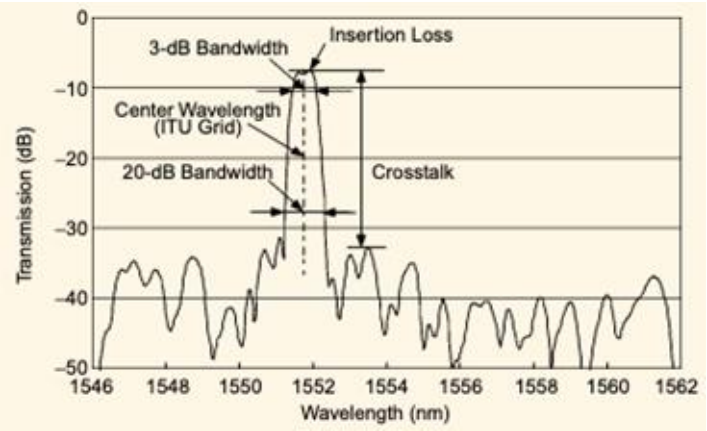
What if we want multi-channel (n x 400G)?

Few things to consider:

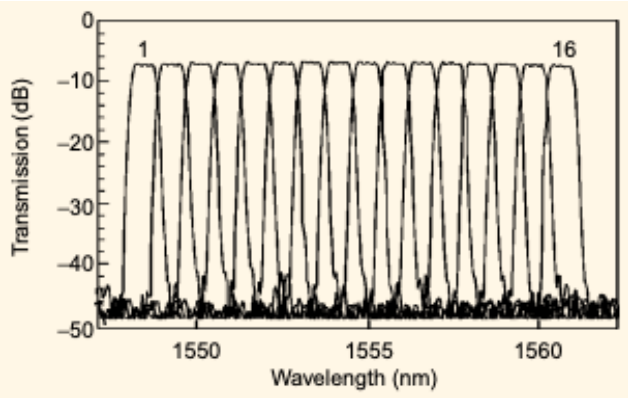
- Mux/deMux:
  - 400G coherent requires 75-100GHz spectrum
    - Does your Mux/deMux support 100GHz or wider/flex-grid slots?
  - Is the filter wide Flat-top or narrow Gaussian passband?
    - Narrow filters could clip 400G signal (60GBd) at the edges

- Insertion loss:
  - Factor in 3-6dB Mux/deMux insertion loss

Narrow-band Gaussian filter
(100GHz spacing)



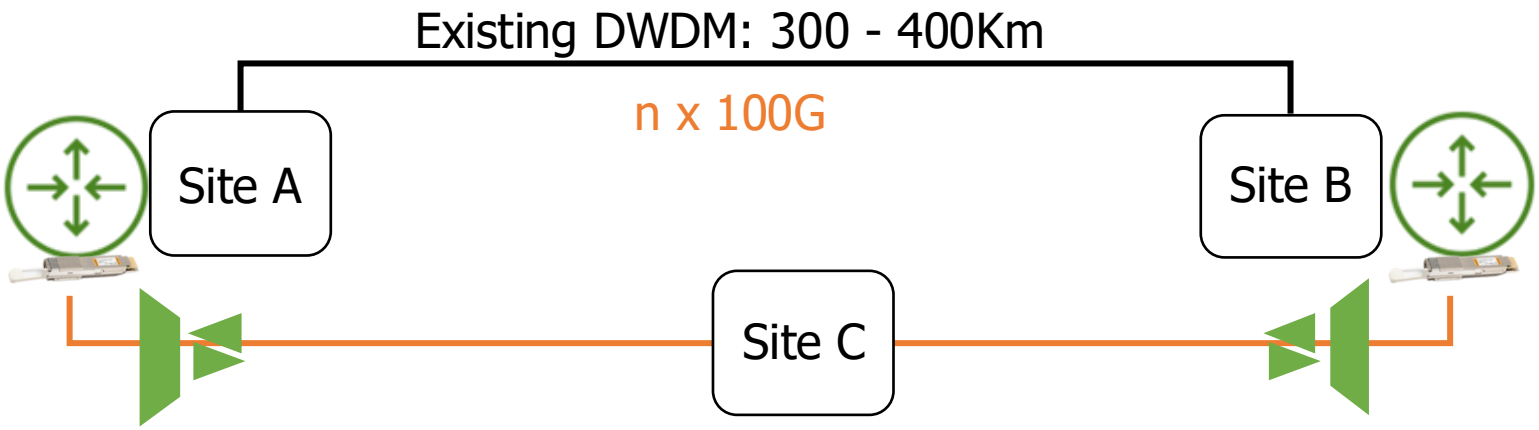
Source: John Hopkins University



Flat-top filter (16 channel)

Use Case3: Long-haul IPoDWDM Links

New build? Or as a backup to your existing DWDM links?



IPoDWDM as backup – with or without drop site
(few x 400G)

Use Case3: Long-haul IPoDWDM Links

Long-haul links require amplification:

- Amplifier restores power (but not signal quality)
  - Also adds ASE noise (NF) ~ noise cannot be removed
  - OSNR declines gradually ~ higher bit errors
- Need to optimise for both RX power and Signal quality (OSNR)
- Proper planning needed:
  - How many inline amplifiers (ILA)? At what span distance? noise figure (NF)?

Use Case3: Long-haul IPoDWDM Links

Note1: Modern coherent plugs allow trading capacity for reach

- Lower OSNR and RX power requirements for lower bit rates
- Can I make do with 4x200G channels instead of 2x400G channels?

Example: 400G ZR+ coherent optics:

Mode	Modulation	RX threshold	OSNR threshold	What it means?	Indicative Reach*
400G ZR+ MSA	DP-16QAM	-12dBm	24 dB	Can handle poorer link conditions: - *Tolerate lower OSNR* - *Longer reach* - *Trade capacity for reach*	~480 Km
300G ZR+	DP-8QAM	-15dBm	21 dB		~600 Km
200G ZR+	DP-QPSK	-18dBm	16 dB		~1000 Km
100G ZR+	DP-QPSK	-18dBm	12.5 dB		~2000 Km

Note2: 400G ZR++ (ULH) with 16.8dB OSNR for 400Gbps already exists!

Use Case4: IPoDWDM with ROADMs

Launch power matters:

- per-channel input into ROADMs (typically -3dBm to 0dBm)
- Will need higher launch to integrate well (400G ZR+: -5dBm to +3dBm)

Alien-wavelength integration:

- Does your line system support alien waves?

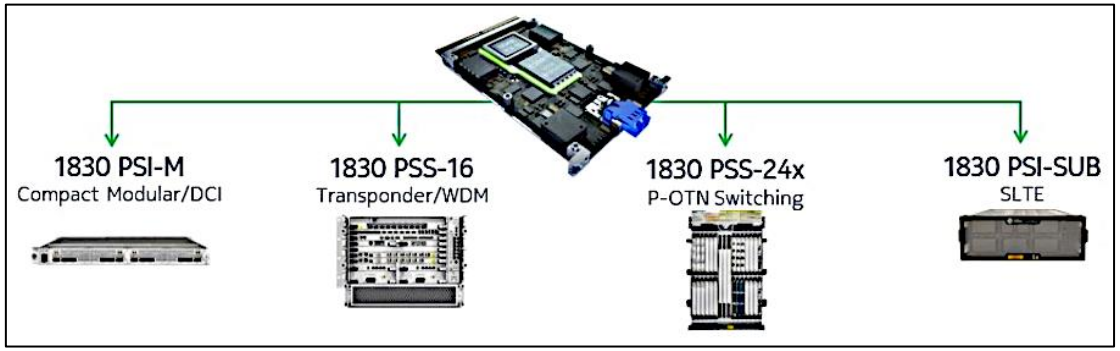
Check route-by-route suitability

- Is every route a pluggable route?
  - For challenging routes, would an embedded optics work better?
  - Is spectral efficiency a concern? Eg: if you must trade capacity for reach?

Embedded optics still have a role

Less constrained by size, power and thermal limits:

- Support higher baud rates
- Can apply more aggressive PCS and FEC optimisation
- Benefits from higher-performance DSPs



Example: Nokia's coherent PSE-6s

Source: Nokia

On more challenging paths:

- (low-OSNR links, long-haul or subsea routes)
- Line cards with embedded optics can offer better *capacity, reach, or margin* than a pluggable

Power and Cooling Considerations

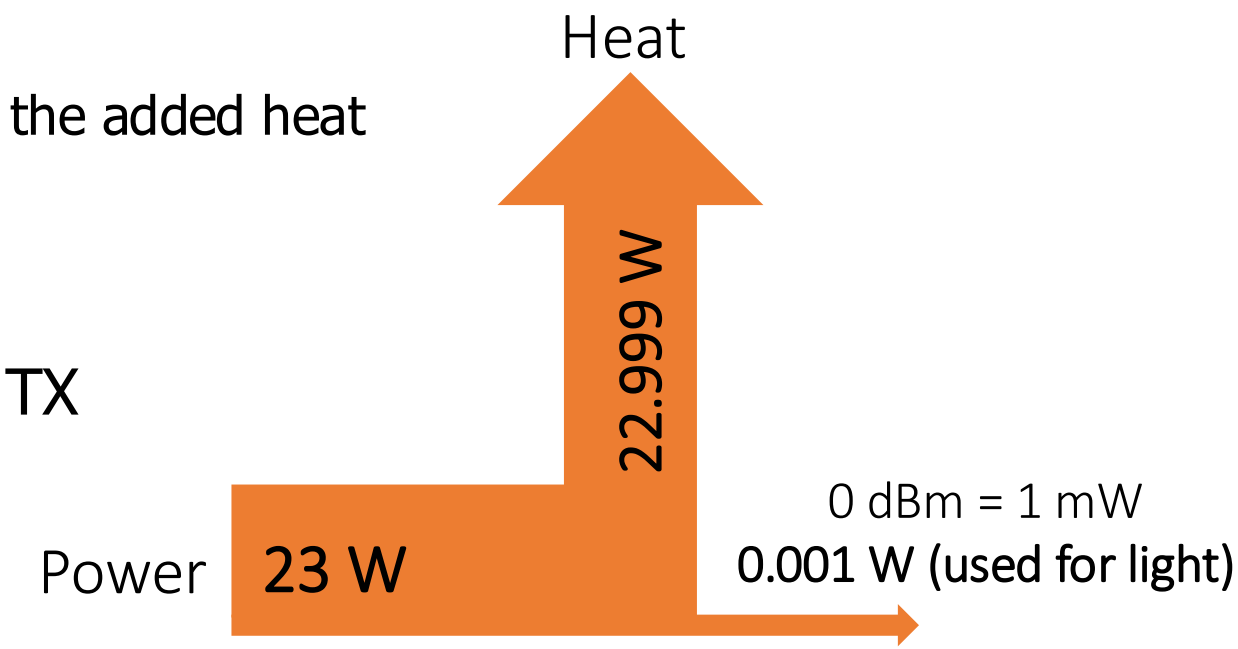
Power: example, 400GZR/ZR+ coherent plugs: 19-23W per port

- Verify PSU headroom
- Check per-port budget, adjacent-port derating and supported coherent population

Cooling checks

- Check rack and room cooling capacity for the added heat

Example: For a 400G ZR+ with 0dBm TX



Other Considerations

Platform checks:

- Validate router/switch support for the coherent plug
  - Form factors: OSPF/QSFP-DD
  - Support for high power plugs: 400G ZR vs 400G ZR+ vs 400G ZR++
- Confirm C-CMIS / application mode support

Network Management:

- Verify the host exposes the required coherent telemetry
- Verify your NMS/Automation tooling can handle coherent telemetry
  - (ingest, visualize and alert on)

```
IF
  OSNR drops below certain threshold
AND
  pre-FEC BER is above certain threshold
AND
  Q-margin drops below certain threshold
THEN
  classify as ”link degradation developing”
AND
 alert NOC and optical team
```

Take aways?

IPoDWDM is no more just a concept

Coherent telemetry now lives on the router

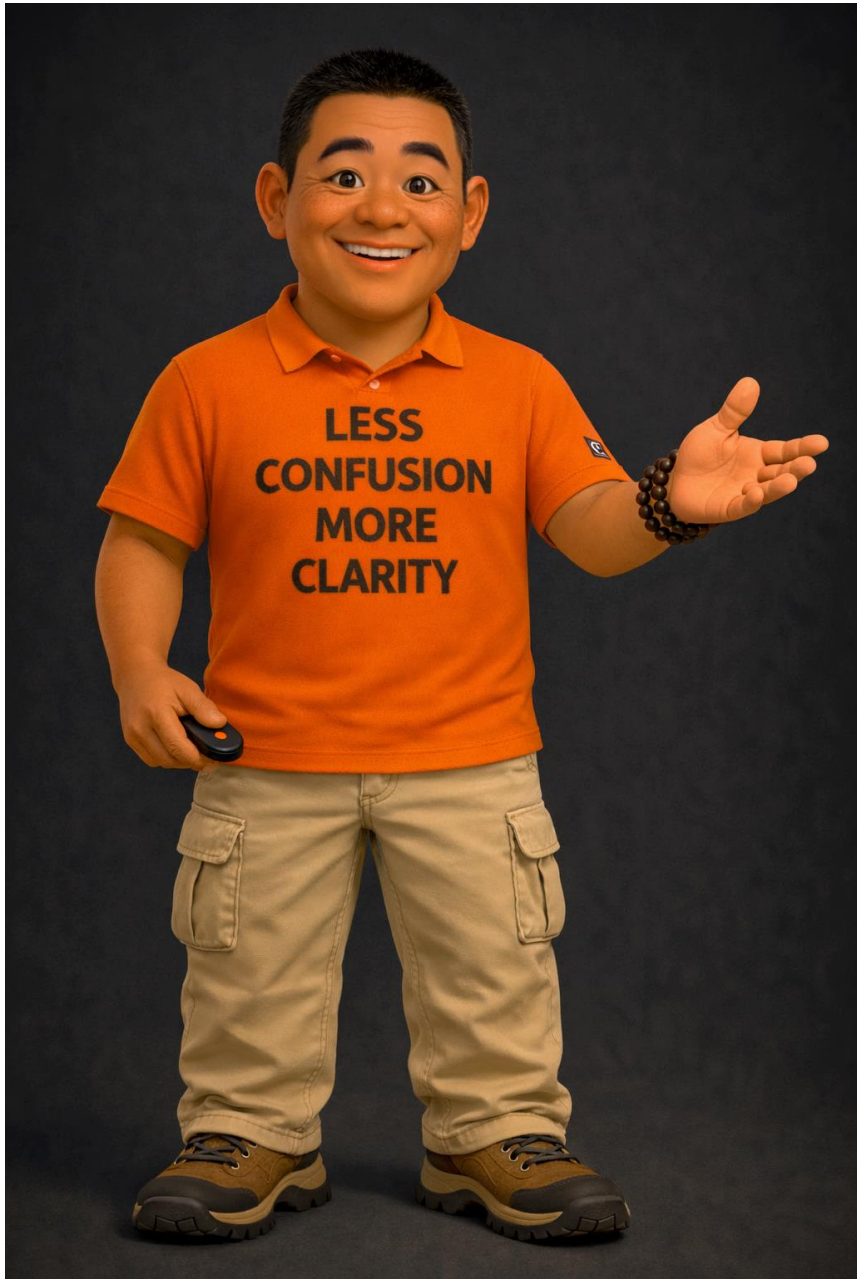
- IP folks will need some understanding of DWDM/coherent concepts

Optical engineering challenges will remain:

- Design and provisioning will require collaboration between IP and optical teams
- Fault isolation will span both IP and optical domains

Disaggregation will continue to expand deployment choice

- Operators can adopt multi-vendor optical components more flexibly


LESS
CONFUSION
MORE
CLARITY