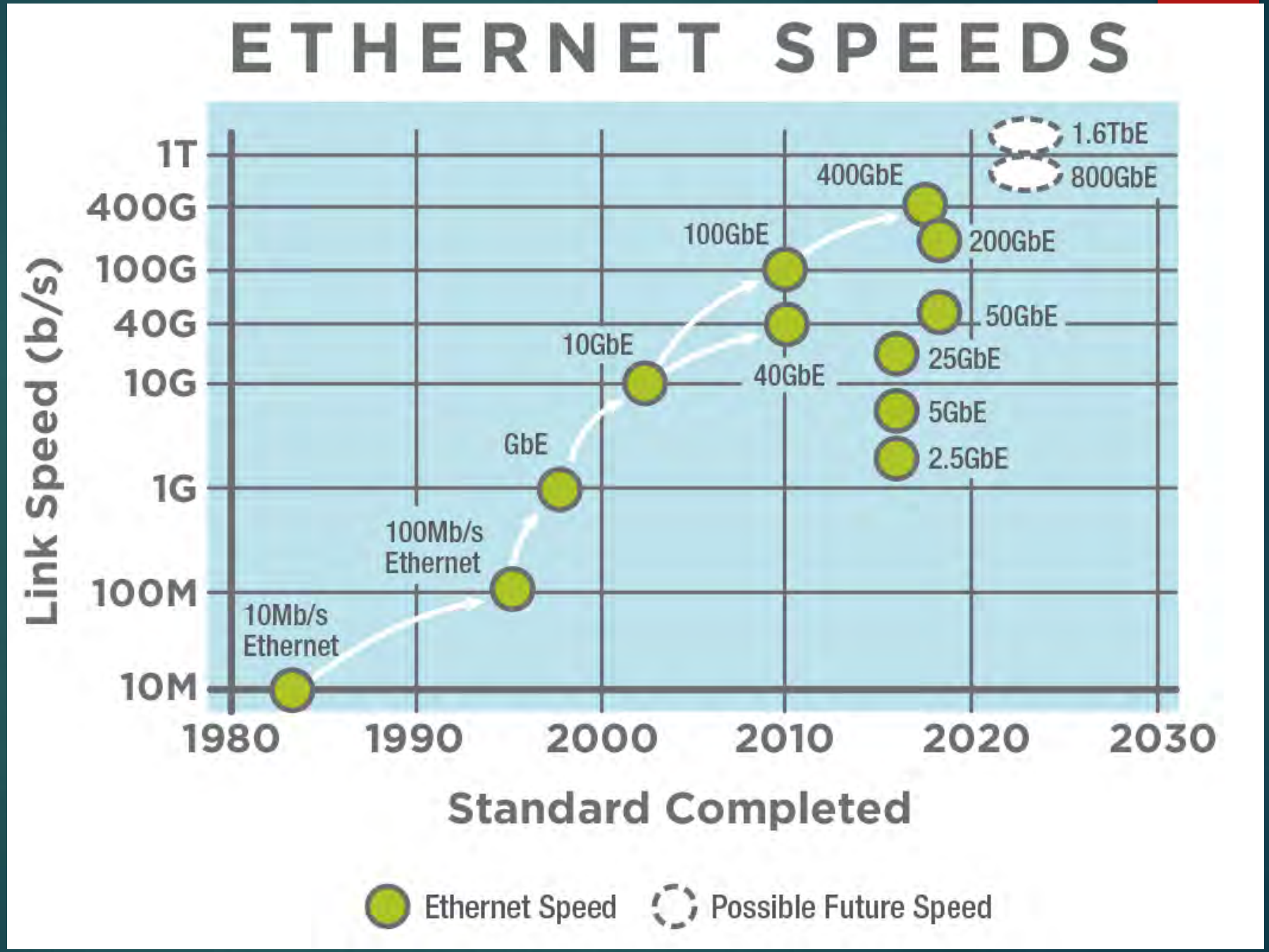
Beyond 25GE

JUG

Topic

- IEEE802.3 standard
- What is FEC?
- FEC clause
- NRZ & PAM4 Signalling
- How to select transceiver module?

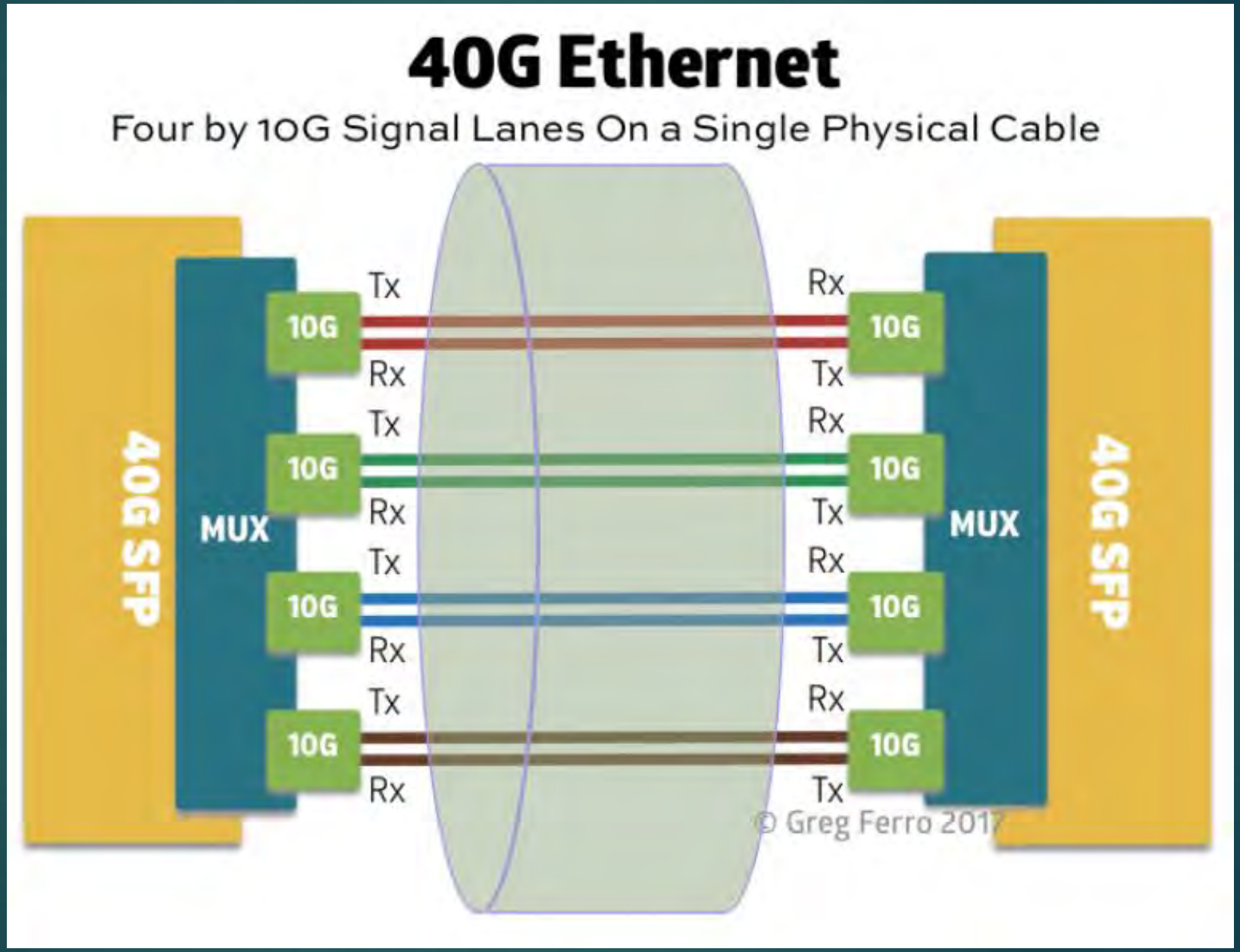
IEEE 802.3 Standard



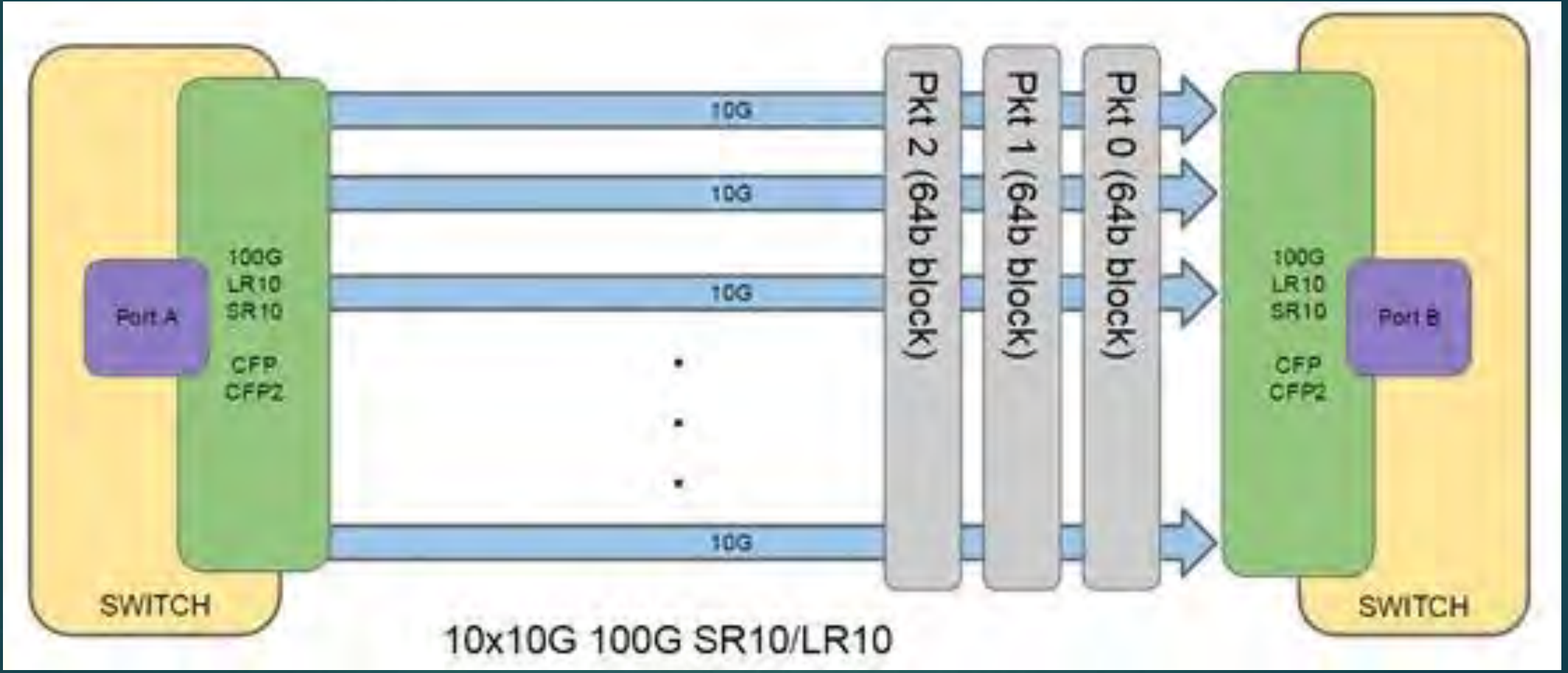
IEEE 802.3 standard

- IEEE 802.3ab (1999) = 1Gbps
- IEEE 802.3ae (2002) = 10Gbps
- IEEE 802.3ba (2010) = 40Gbps & 100Gbps
- IEEE 802.3by (2016) = 25Gbps single lane and 50Gbps dual lane
- IEEE 802.3bz (2016) = 2.5GBASE-T and 5GBASE-T (NBASE-T)
- IEEE 802.3bs (2017) = 400Gbps
- IEEE 802.3cd (2018) = 50Gbps single lane and 200Gbps

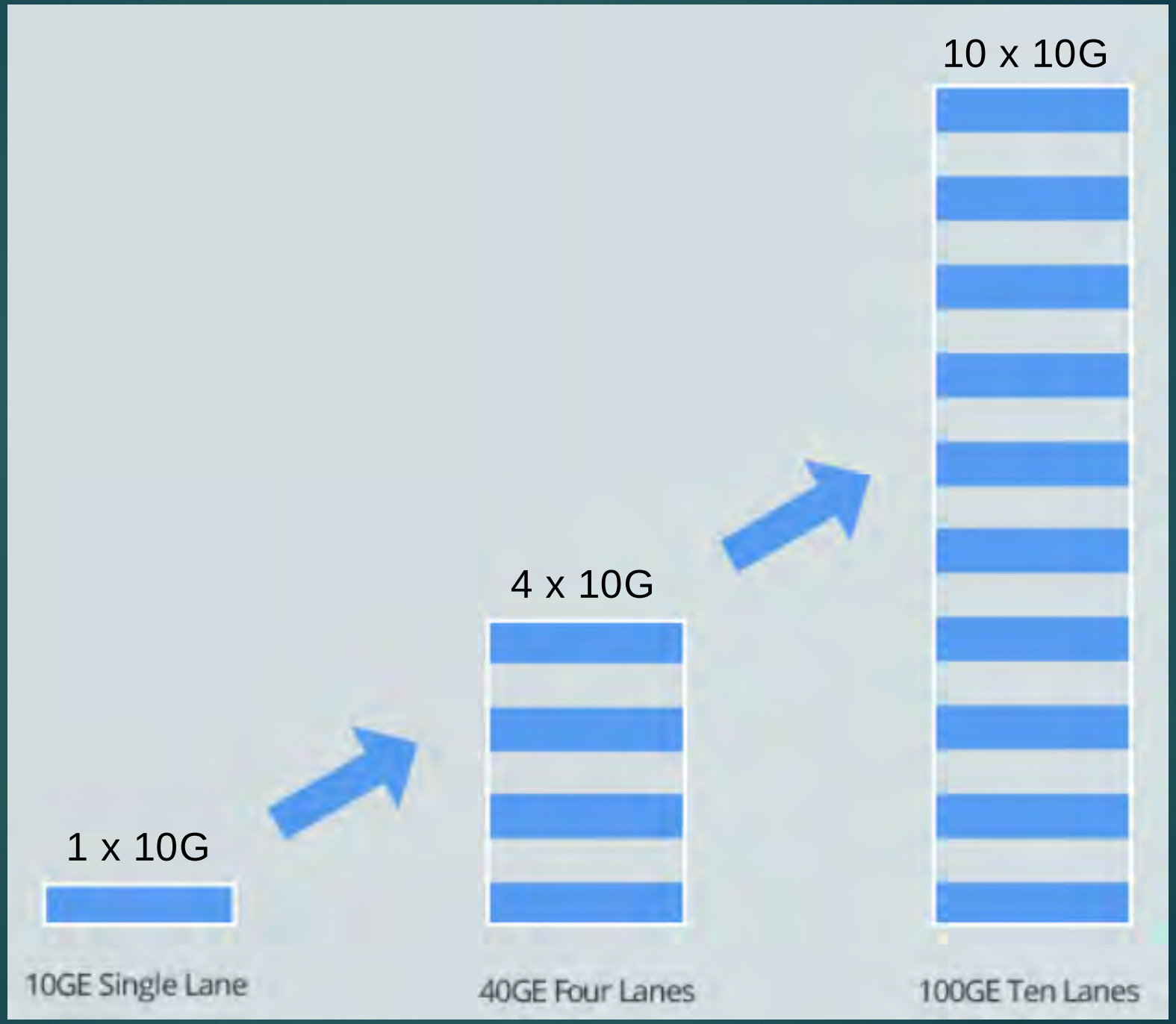
What is lane?


40G Ethernet
Four by 10G Signal Lanes On a Single Physical Cable
40G SFP
MUX
10G
Tx
Rx
10G
Tx
Rx
10G
Tx
Rx
10G
Tx
Rx
Rx
Tx
10G
Rx
Tx
10G
Rx
Tx
10G
Rx
Tx
10G
MUX
40G SFP
© Greg Ferro 2017

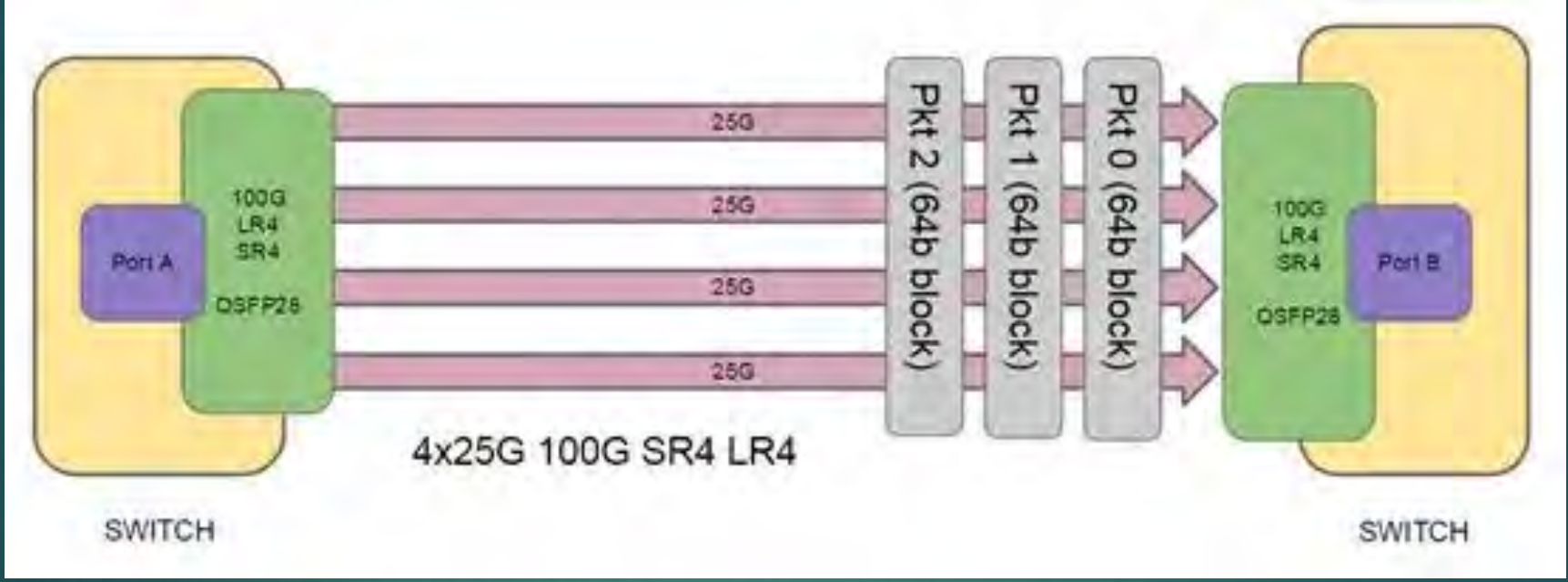
100G (10 x 10G lanes)

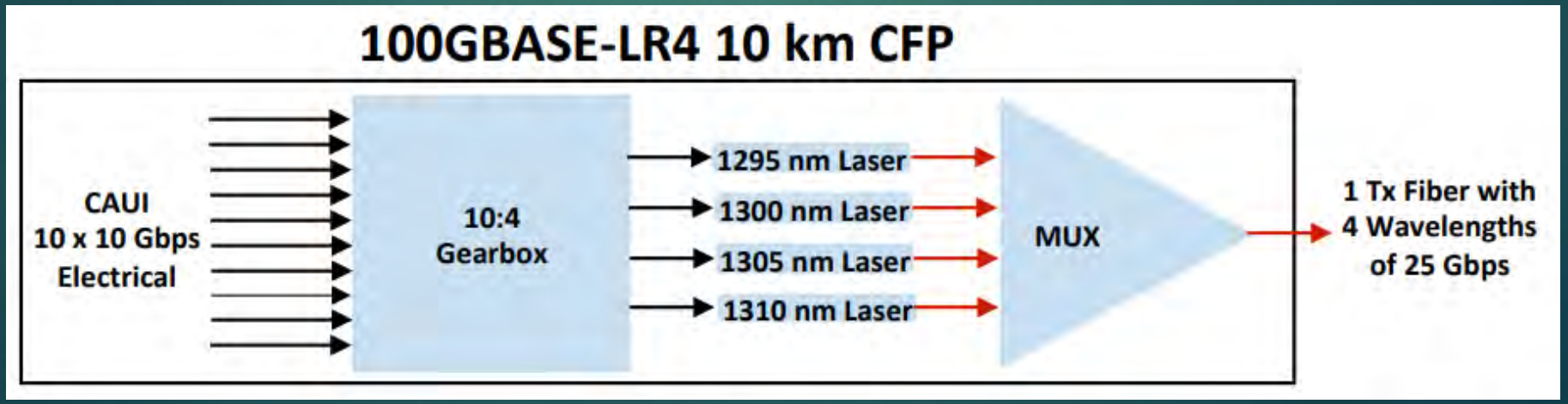


High Speed Ethernet Upgrade Path

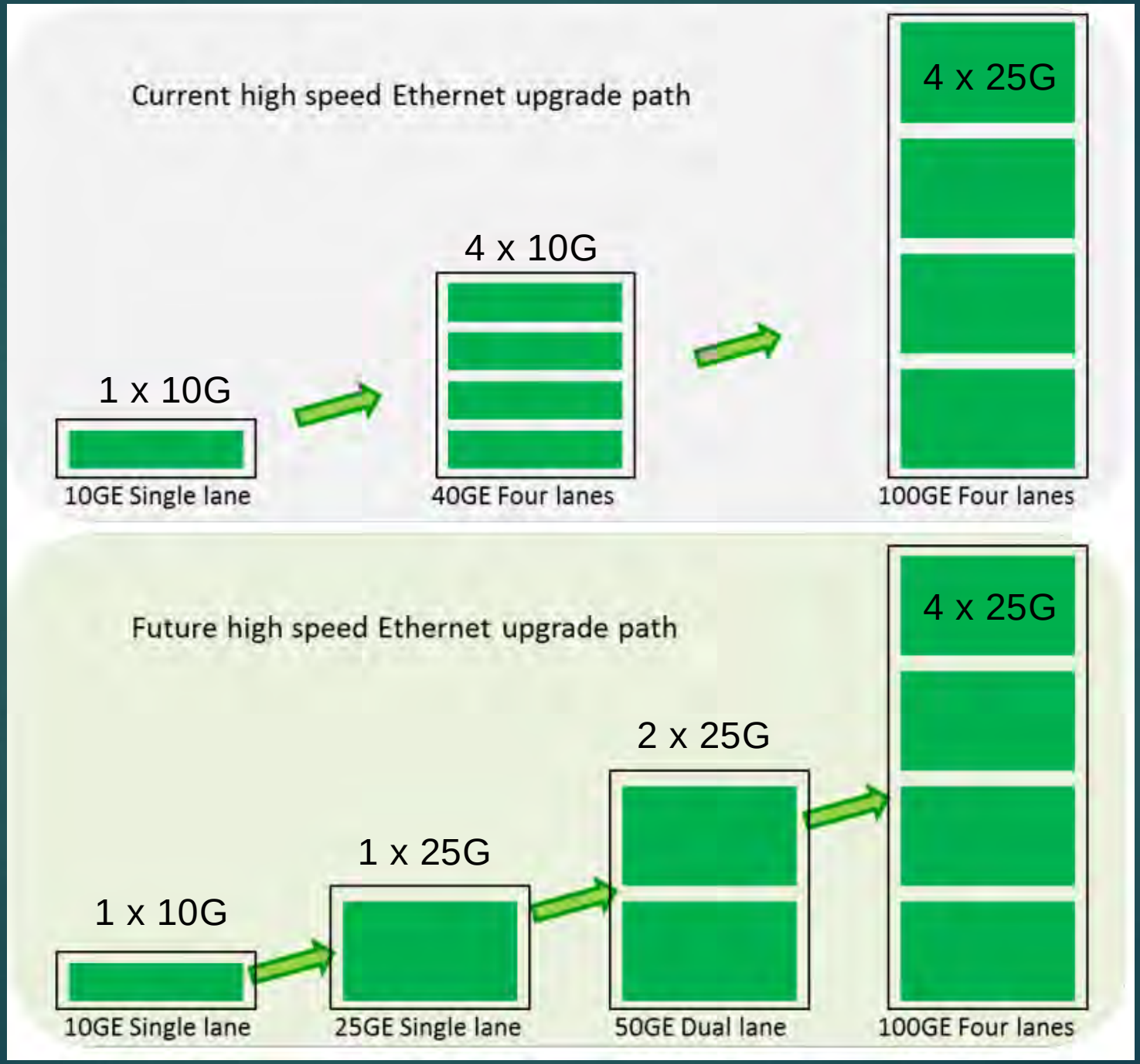


100G (4 x 25G lanes)

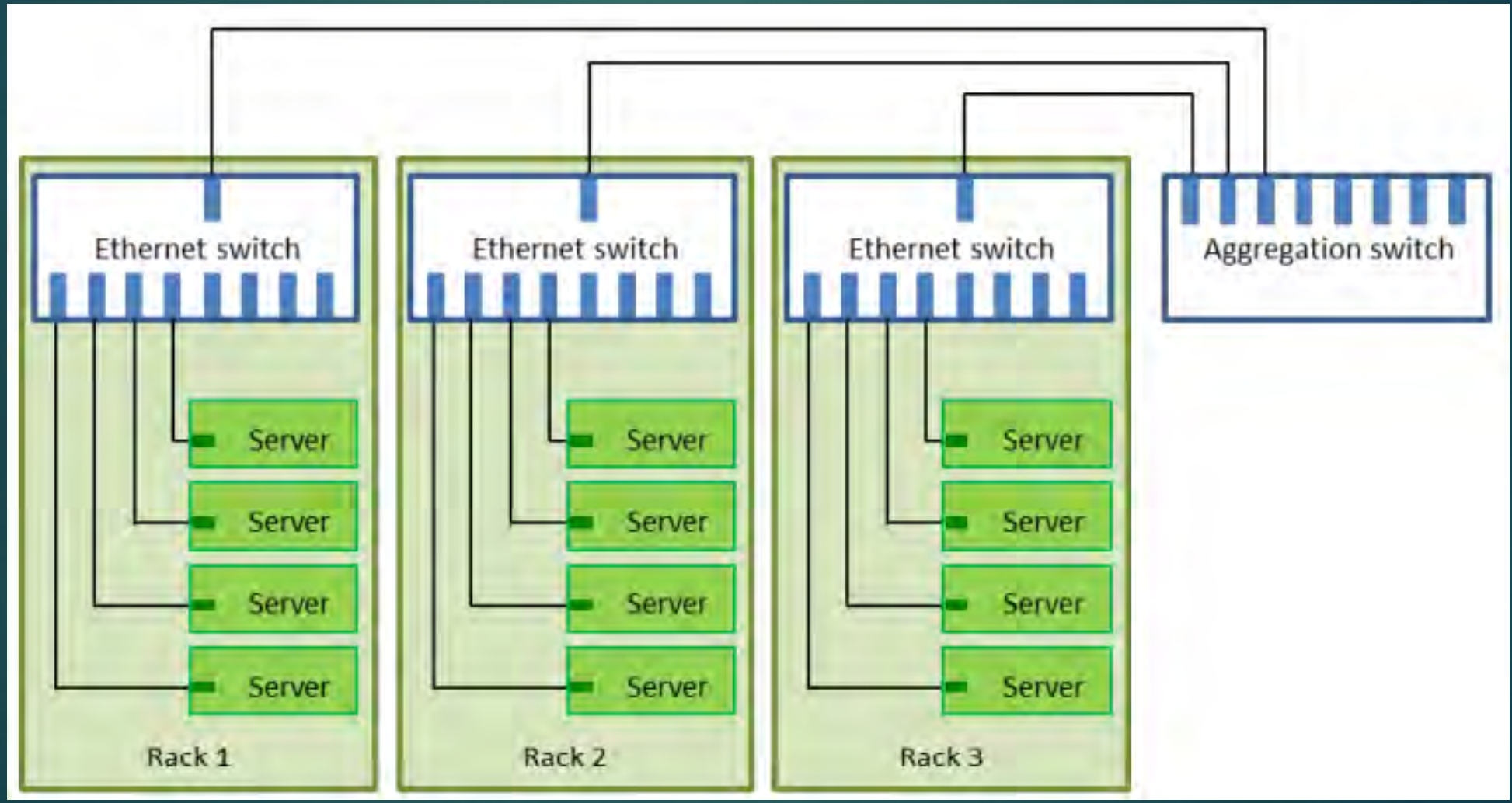




Why need 25G?


Current high speed Ethernet upgrade path
1 x 10G
10GE Single lane
4 x 10G
40GE Four lanes
4 x 25G
100GE Four lanes
Future high speed Ethernet upgrade path
1 x 10G
10GE Single lane
1 x 25G
25GE Single lane
2 x 25G
50GE Dual lane
4 x 25G
100GE Four lanes

Top-of-Rack (ToR) switch design



ToR Design

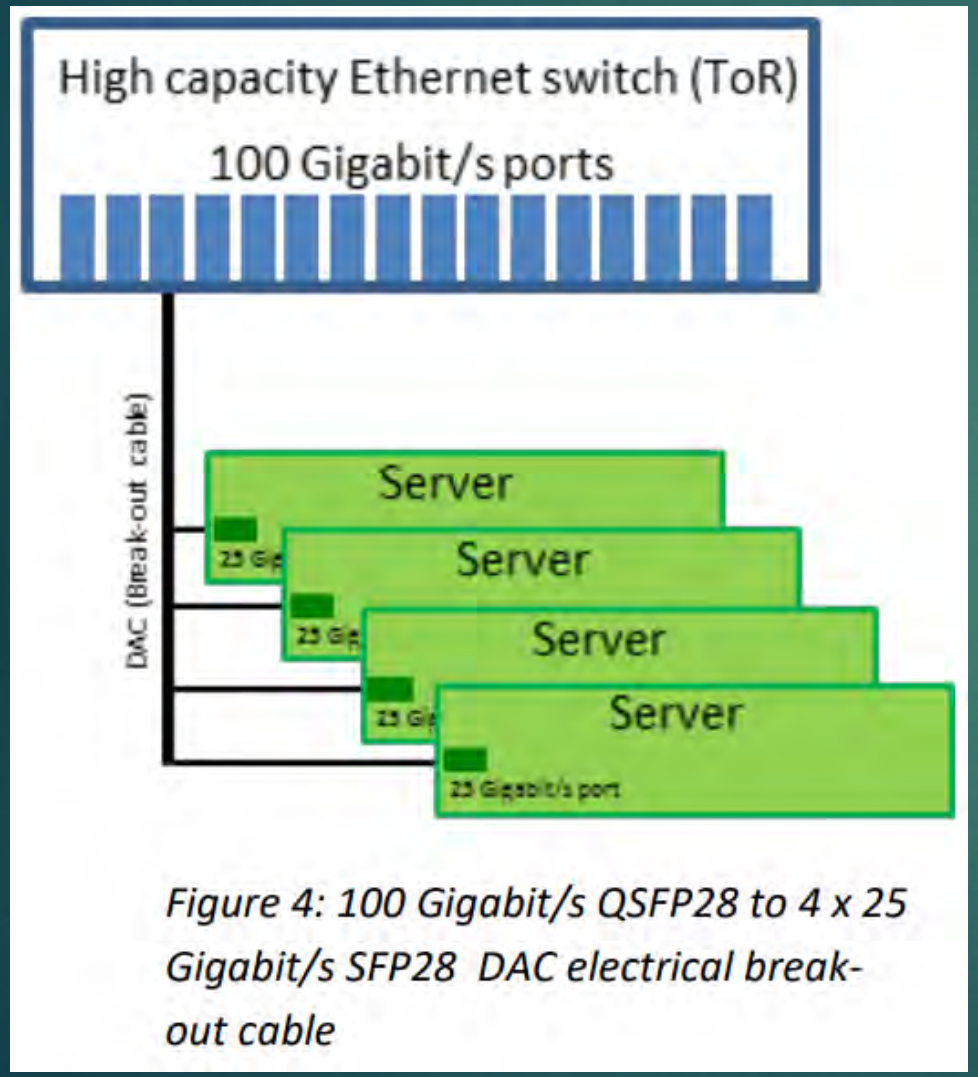
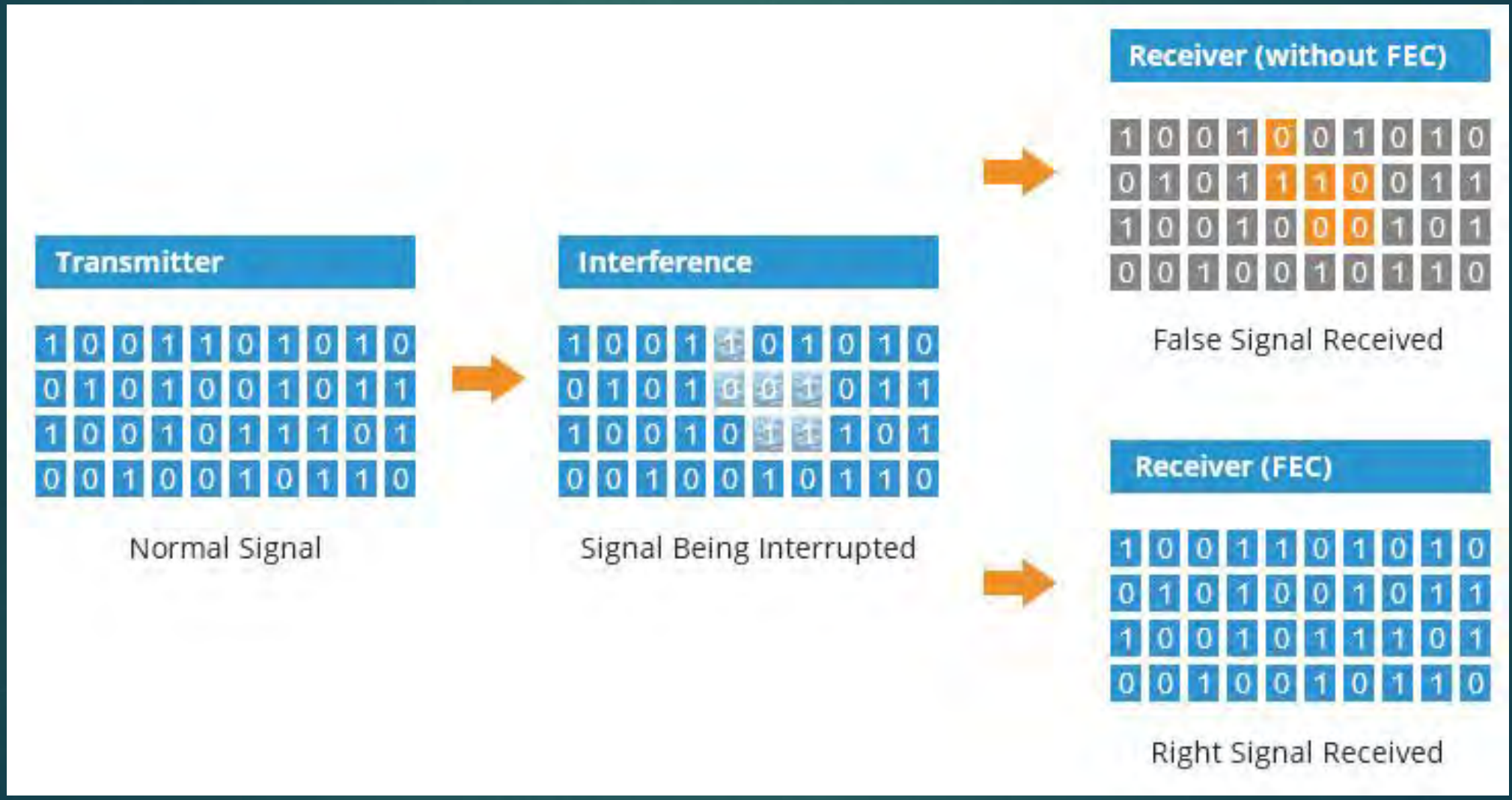
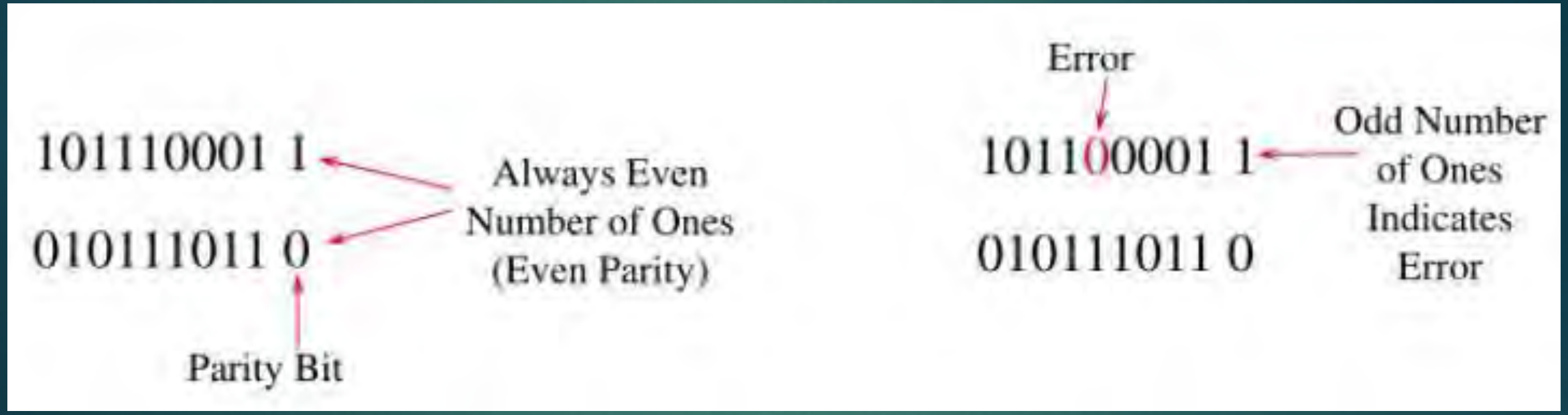


Figure 4: 100 Gigabit/s QSFP28 to 4 x 25 Gigabit/s SFP28 DAC electrical break-out cable

What is FEC? (Forward Error Correction)

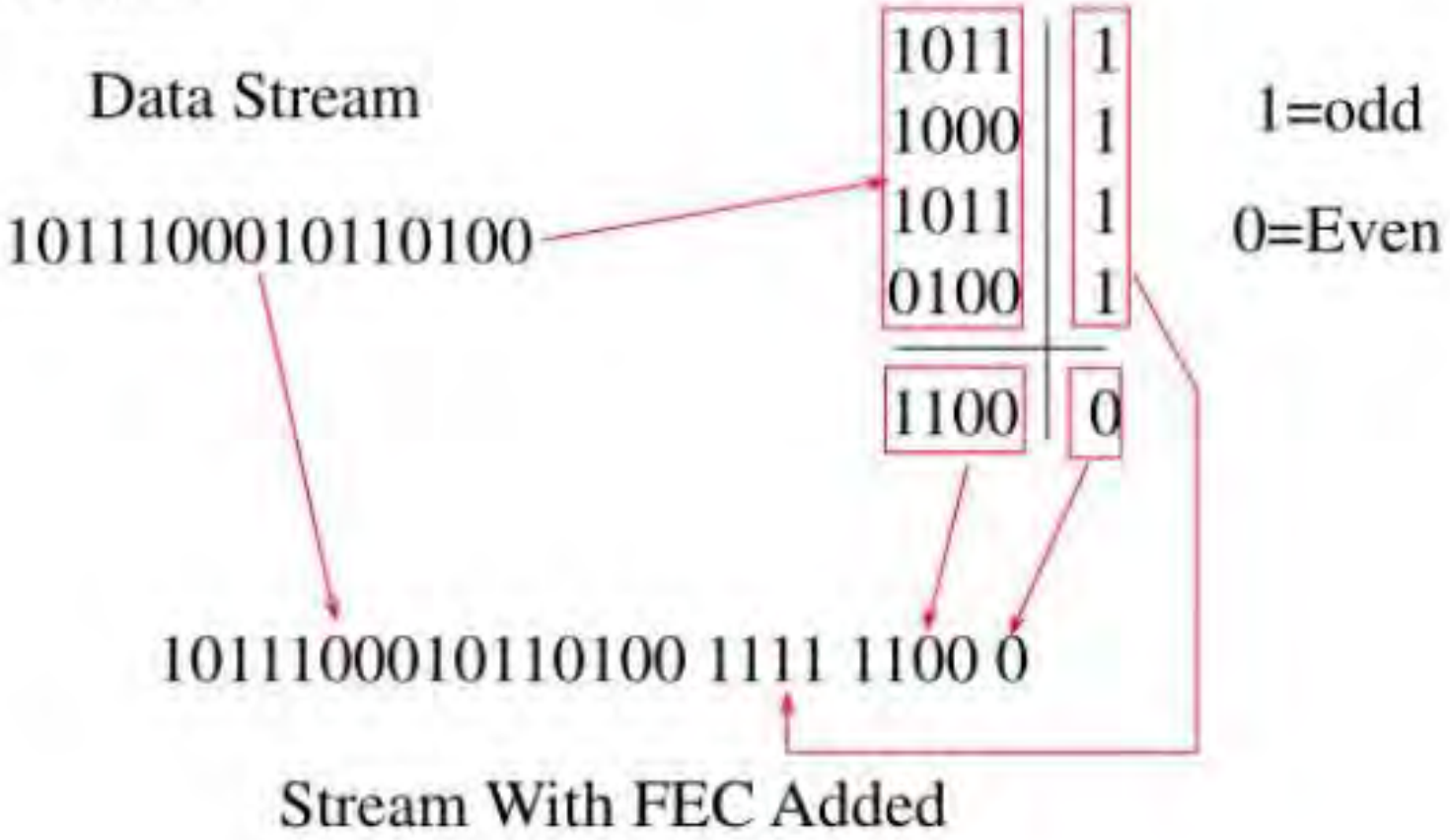


How to Detect Bit Error



How FEC Works

- FEC works by the addition of additional data bits to the data stream to determine if errors exist and to try and correct them.

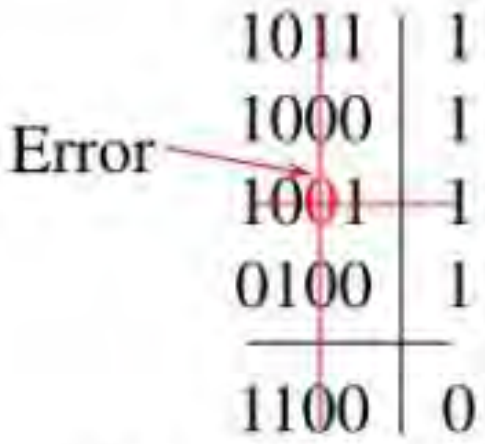


How FEC Works

- Once you know a bit is wrong, correcting it is easy, if you know its wrong and its a zero, then it has to be a one.

1011	1
1000	1
1011	1
0100	1
1100	0

Before Transmission



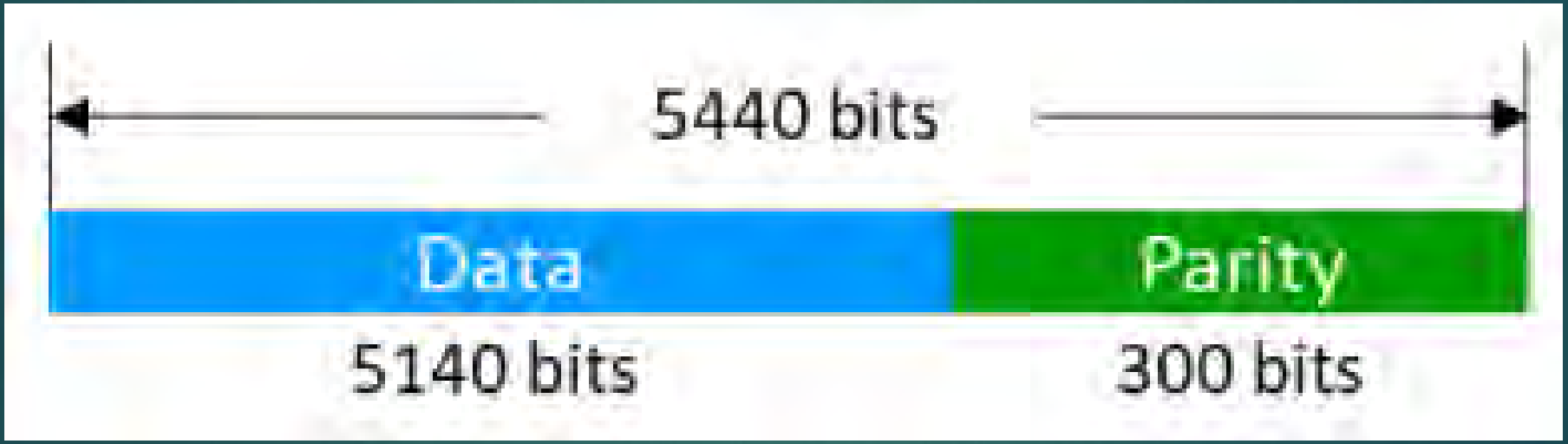
After Transmission the Bit in Error is Detected and Corrected

1=odd

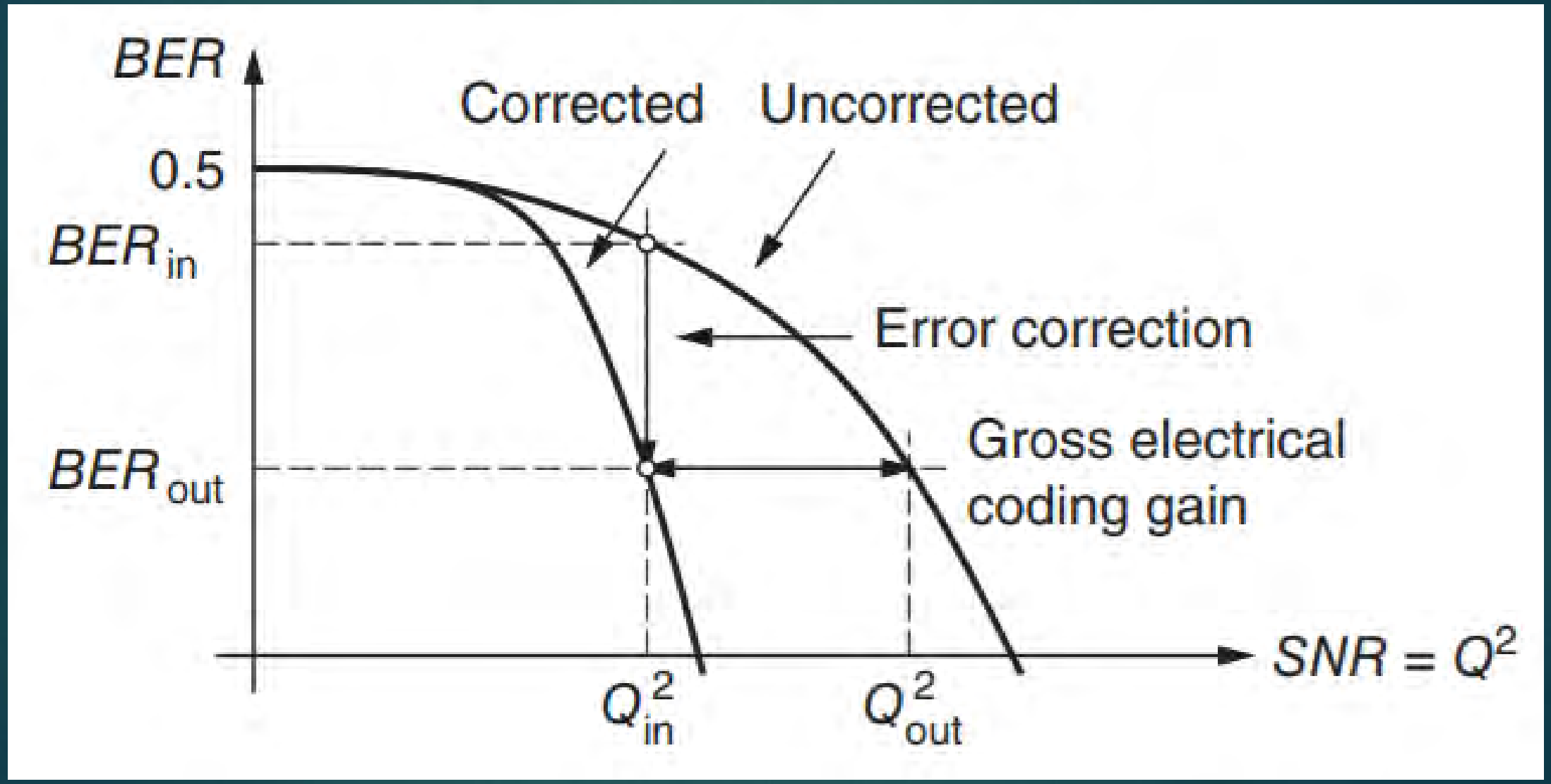
0=Even

Parity Bit

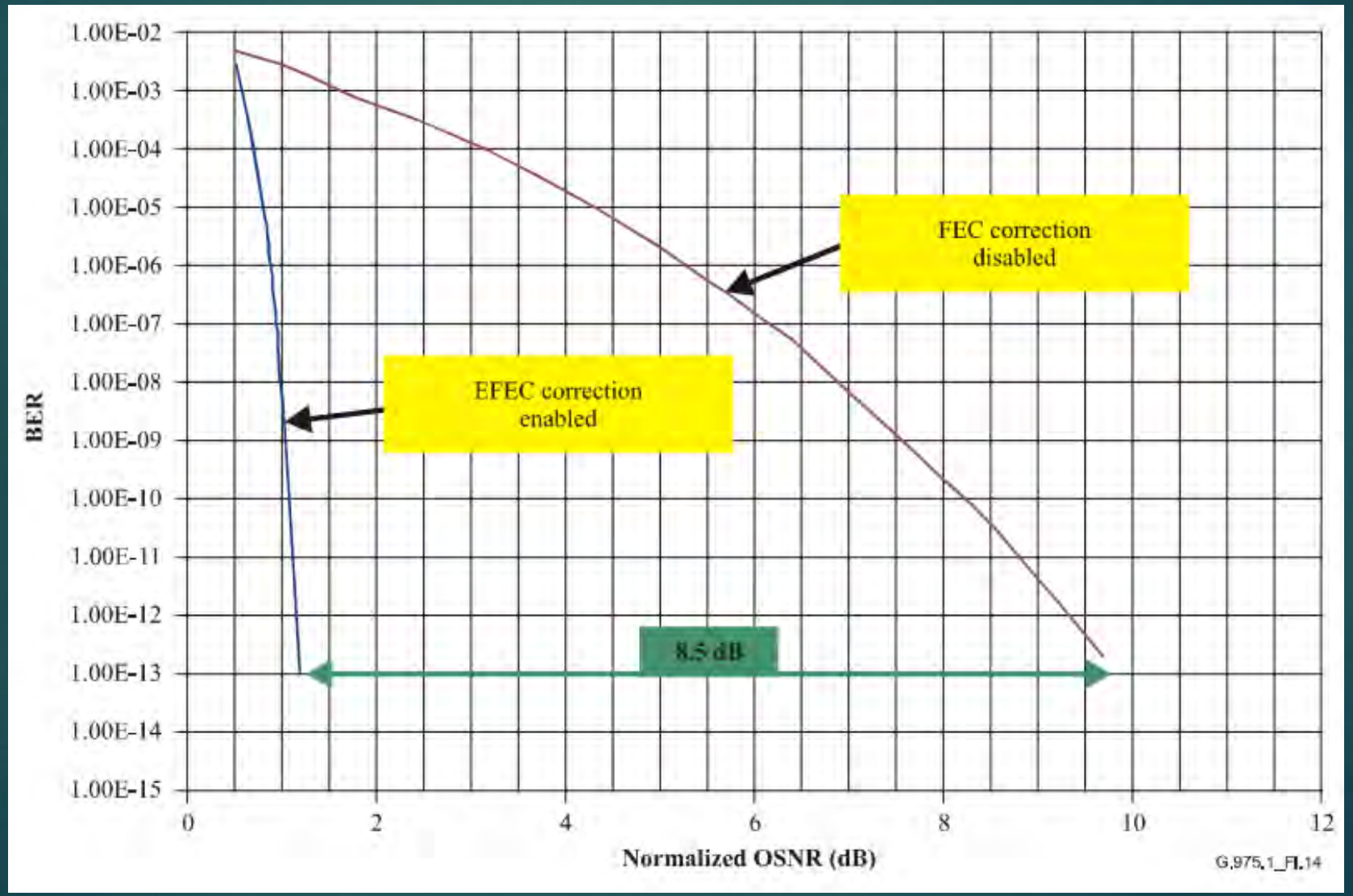
- RS(5280,5140)
- RS(528,514)



Coding Gain


BER
Corrected
Uncorrected
0.5
BER_{in}
Error correction
Gross electrical
coding gain
BER_{out}
$SNR = Q^2$
Q^2_{in}
Q^2_{out}

Coding Gain


1.00E-02
1.00E-03
1.00E-04
1.00E-05
1.00E-06
1.00E-07
1.00E-08
1.00E-09
1.00E-10
1.00E-11
1.00E-12
1.00E-13
1.00E-14
1.00E-15
BER
0
2
4
6
8
10
12
Normalized OSNR (dB)
FEC correction disabled
EFEC correction enabled
8.5 dB
G.975.1_FI.14

FC-FEC (IEEE Clause 74)

- Name Fire code FEC, FC-FEC, BASE-R FEC, CL74
- Parity 32 bits (2112, 2080)
- Suitable for 25GE (NRZ signalling)
- Mainly aimed at bursts due to DEF error propagation
  Random error correction capability is not strong
- Estimate input Pre-BER : < 1e-8
- Coding Gain : 2.0 – 2.5 dB

RS-FEC (IEEE Clause 91)

- Name Reed-Solomon FEC, RS-FEC, CL91
- Parity 140 bits RS(5280,5140), RS(528,514)
- Suitable for 100GE (NRZ Signalling)
- Random error correction capability is strong
- Estimate input BER < 1e-5
- Coding Gain 5.0 – 5.5 dB

RS-FEC (IEEE Clause 108)

- Name Reed-Solomon FEC, RS-FEC, CL108
- Parity 140 bits RS(5280,5140), RS(528,514)
- Suitable for 25GE (NRZ Signalling)

FEC and Latency

FEC	Latency (ns)	BER
None	0	1e-12
BASE-R FEC 1 lane (25 Gb/s)	82	1e-8
RS-FEC 1 lane (25 Gb/s)	250	1e-5

25GE FEC Requirement (Arista Networks)

Table 3: 25G Optics and Cables from Arista and Minimum FEC Requirements		
SKU	**Description**	**Minimum FEC Requirement**
CAB-S-S-25G-xM	25G SFP to SFP twinax copper cable, 1, 2, 3 meters	No FEC (CA-N type)
CAB-Q-4S-100G-xM	100GBASE-CR4 QSFP to 4 x 25GbE SFP Twinax Copper Cable, 1, 2 meter	No FEC (CA-N type)
CAB-Q-4S-100G-3M	100GBASE-CR4 QSFP to 4 x 25GbE SFP Twinax Copper Cable, 3 meter	BASE-R FEC (CA-S type)
SFP-25G-SR	25GBASE-SR SFP transceiver up to 70m over parallel OM3 or 100m over OM4 multimode fiber	RS-FEC
SFP-25G-LR	25GBASE-LR SFP transceiver, up to 10km over single-mode fiber	RS-FEC
AOC-S-S-25G-xM	25GbE SFP to SFP Active Optical Cable, 3m to 30m lengths	RS-FEC

25GE FEC Requirement (Cisco)

SFP Type and Distance	FEC Type
25G Copper Cable (DAC) ≤ 2m	No FEC
25G Copper Cable (DAC) 2m < X ≤ 3m	BASE-R FEC
25G Copper Cable (DAC) > 3m	RS-FEC
25G Active Optical Cable (AOC)	BASE-R FEC or RS-FEC
25G-SR (OM3 = 70m / OM4 = 100m)	RS-FEC
25G-BR (SMF) ≤ 1.5 Km	No FEC
25G-BR (SMF) ≤ 3.5 Km	BASE-R FEC
25G-BR (SMF) > 3.5 Km	RS-FEC

Refer :https://www.cisco.com/c/en/us/products/collateral/interfaces-modules/transceiver-modules/datasheet-c78-736950.html

25GE FEC for Dell EMC

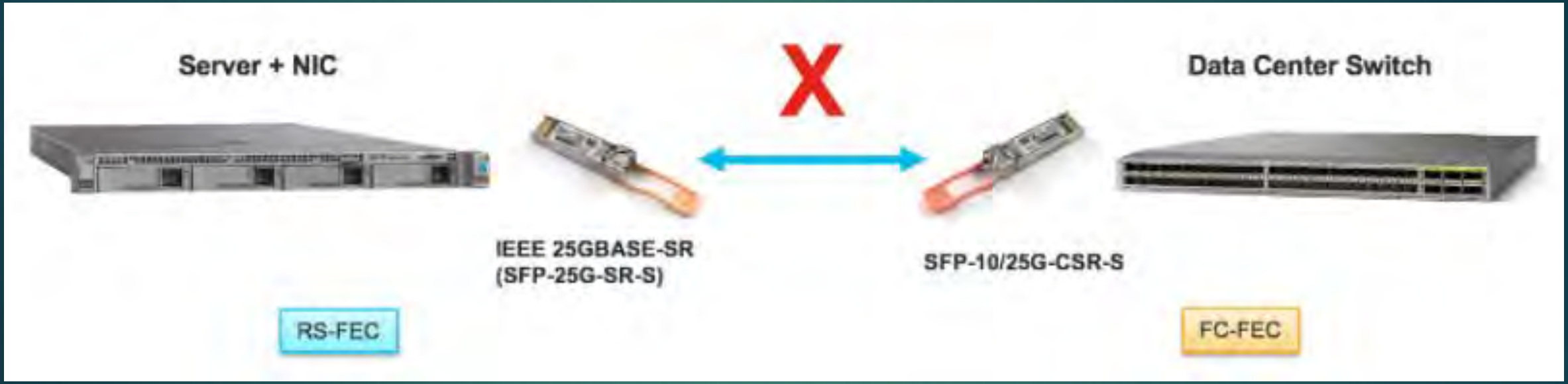
Optic/Cable	SKU	Dell P/N	AutoNeg (Default)	FEC (Default)	FEC Clause (CL) type Enabled
SFP28 25G DAC 1m	470-ACES	2JVDD	Enabled	Not requested	
SFP28 25G DAC 2m	470-ACET	DOR73	Enabled	Not requested	
SFP28 25G DAC 3m	470-ACEV	VXFJY	Enabled	Not requested	
SFP28 25G DAC 5m	470-ACEW	9X8JP	Enabled	Requested	CL108
SFP28 25G AOC 7m	470-ACIM	3YWG7	Disabled	Enabled	CL108
SFP28 25G AOC 10m	470-ACIN	5CMT2	Disabled	Enabled	CL108
SFP28 25G AOC 15m	470-ACIL	RCVP5	Disabled	Enabled	CL108
SFP28 25G AOC 20m	470-ACIG	X5DH4	Disabled	Enabled	CL108
100G to 4*25G DAC 1m	470-ABOR	26FN3	Enabled	Not requested	
100G to 4*25G DAC 2m	470-ABOS	YFNDD	Enabled	Not requested	
100G to 4*25G DAC 3m	470-ABOT	7R9N9	Enabled	Requested	CL108
SFP28 SR	407-BBWO	P7D7R	Disabled	Enabled	CL108
SFP28 SR-NOF	407-BBXX	W4GPP	Disabled	Enabled	CL108
SFP28 LR	407-BBXY	0YR96	Disabled	Disabled	

FECs match

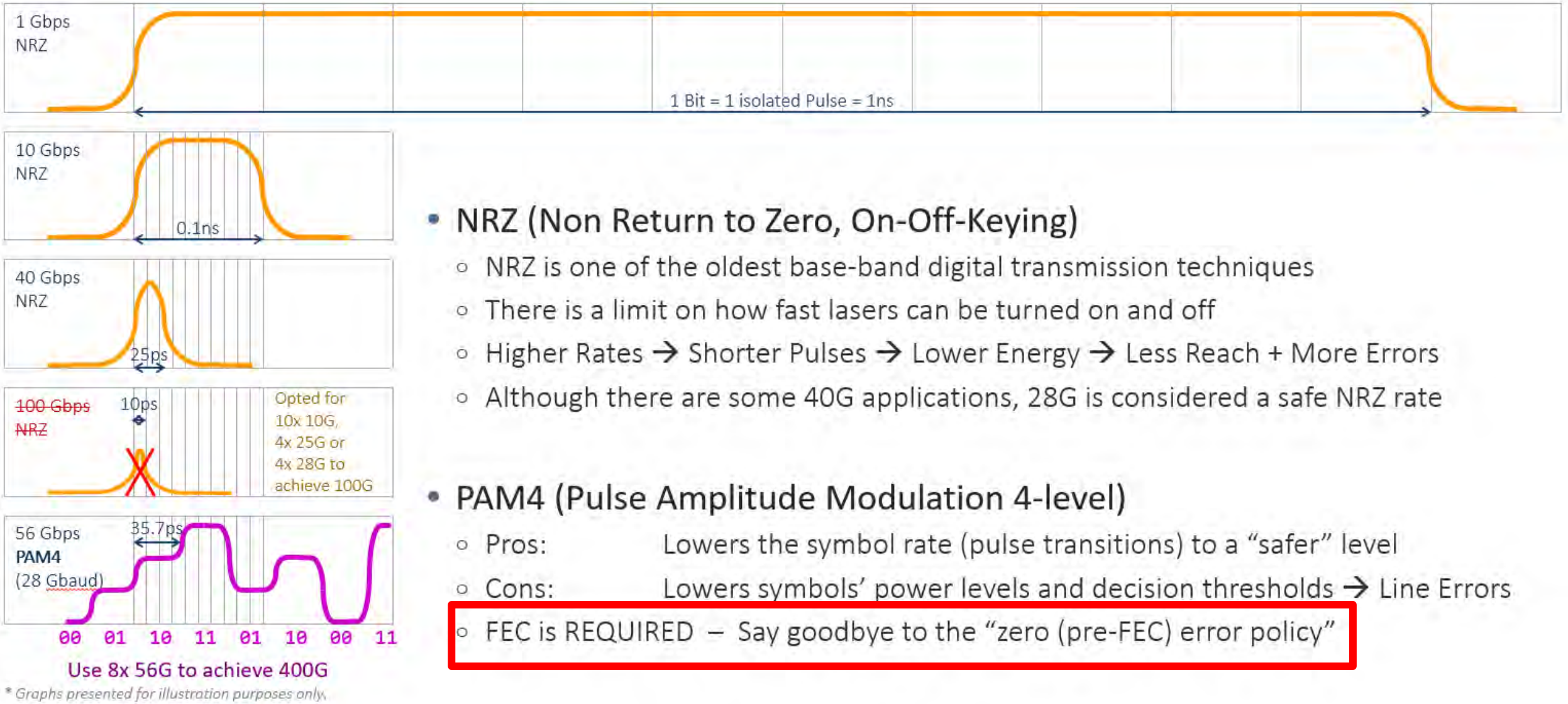

Server + NIC
Data Center Switch
SFP-10/25G-CSR-S
SFP-10/25G-CSR-S
FC-FEC
FC-FEC


Aggregation switch
Access switch
IEEE 25GBASE-SR
(SFP-25G-SR-S)
SFP-10/25G-CSR-S
RS-FEC
RS-FEC

FECs match


Server + NIC
X
Data Center Switch
IEEE 25GBASE-SR
(SFP-25G-SR-S)
SFP-10/25G-CSR-S
RS-FEC
FC-FEC

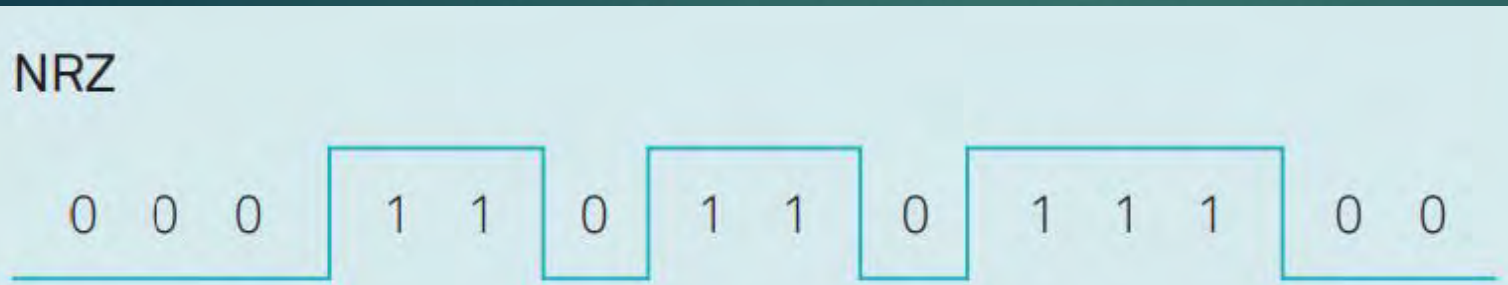
NRZ and PAM4 signaling

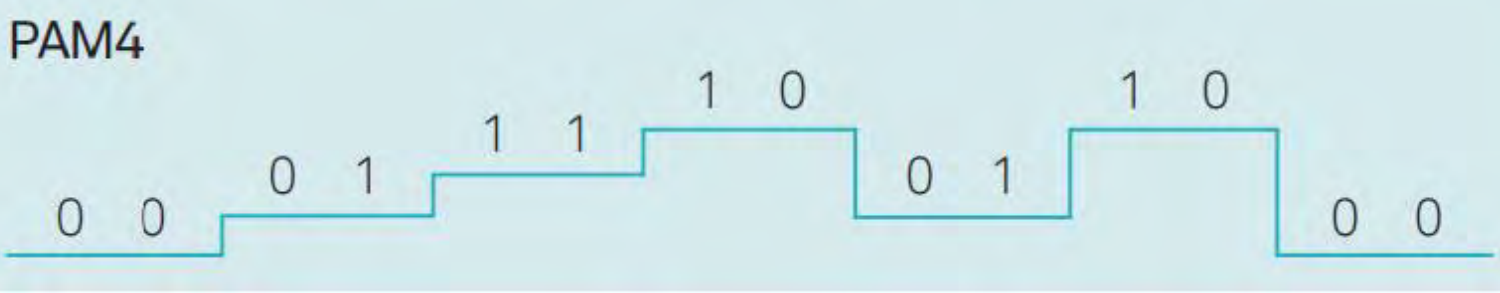


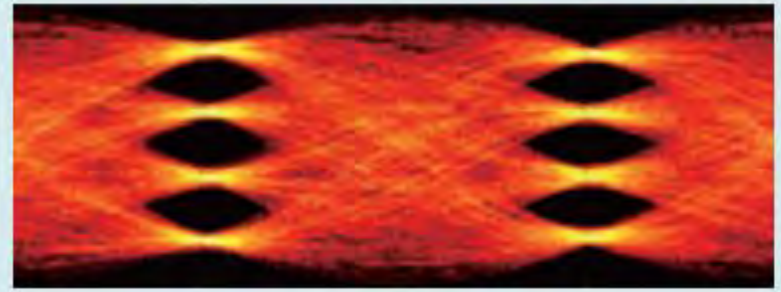
- NRZ (Non Return to Zero, On-Off-Keying)
  - NRZ is one of the oldest base-band digital transmission techniques
  - There is a limit on how fast lasers can be turned on and off
  - Higher Rates → Shorter Pulses → Lower Energy → Less Reach + More Errors
  - Although there are some 40G applications, 28G is considered a safe NRZ rate

- PAM4 (Pulse Amplitude Modulation 4-level)
  - Pros: Lowers the symbol rate (pulse transitions) to a "safer" level
  - Cons: Lowers symbols' power levels and decision thresholds → Line Errors
  - FEC is REQUIRED – Say goodbye to the "zero (pre-FEC) error policy"

NRZ and PAM4 signaling







PAM4 optical transceiver

Interface	Bandwidth	Electrical I/O	Optical I/O	Technology
50GBASE-LR/ER	50 Gbit/s	2 x 25 Gbit/s NRZ	1 x 50G PAM4	1 x 50 Gbit/s 1310 nm PAM4, 1λ
200GBASE-LR4	200 Gbit/s	8 x 25 Gbit/s NRZ	4 x 50G PAM4	4 x 50G PAM4 LAN-WDM, 4λ
400GBASE-LR8	400 Gbit/s	16 x 25 Gbit/s NRZ	8 x 50G PAM4	8 x 50G PAM4 LAN-WDM, 8λ

RS-FEC (IEEE Clause 134)

- Name Reed-Solomon FEC, KP4 FEC
- Parity 300 bits RS(5440,5140), RS(544,514)
- Suitable for 50GE (PAM4 Signalling)
- Random error correction capability is strong
- Coding Gain 6.0 – 6.5 dB

How to select transceiver module?

- Media Type (DAC, AOC, MMF, SMF, CWDM, DWDM)
- Distance (5m, 100m, 2km, 10km, 40km, 80km)
- Connector type (LC, Bidi, MPO)
- Data Lane (1x10G, 4x10G, 1x25G, 2x25G, 1x50G, 4x25G, 2x50G)
- NRZ or PAM4
- FEC type
- Form factor (SFP+, SFP28, QSFP+, QSFP28, SFP56, QSFP56)

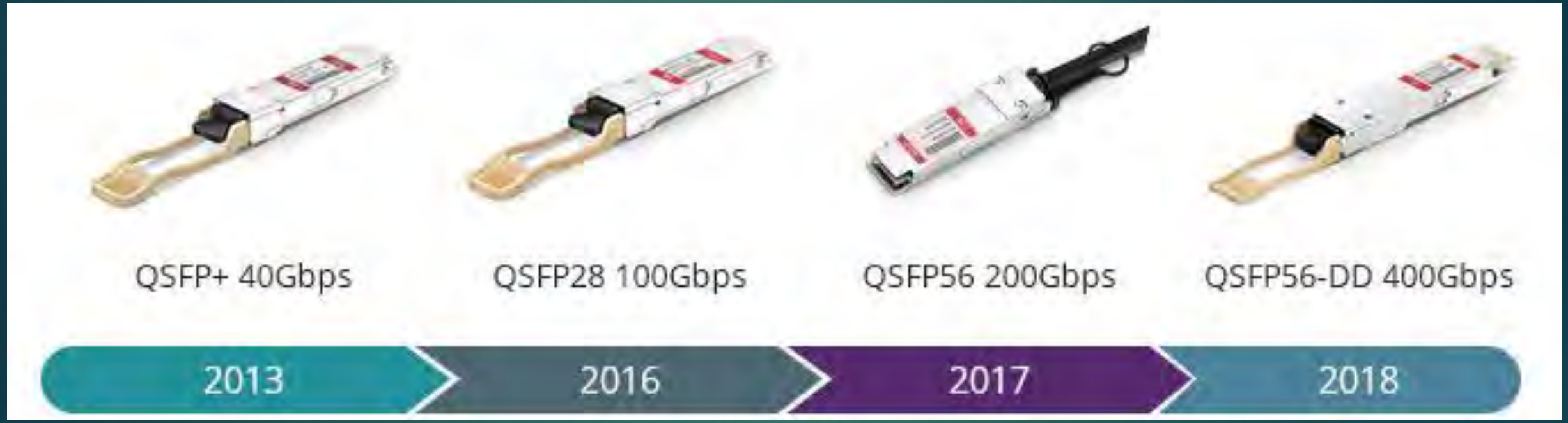
Transceiver Form

SFP = Small Form Pluggable
QSFP = Quad Small Form Pluggable

1 x 1G 1 x 10G 1 x 25G 4 x 10G 4 x 25G

Transceiver Form

QSFP-DD = QSFP Double Density



4 x 10G	4 x 25G	4 x 50G (PAM4)	8 x 50G (PAM4)

Q & A