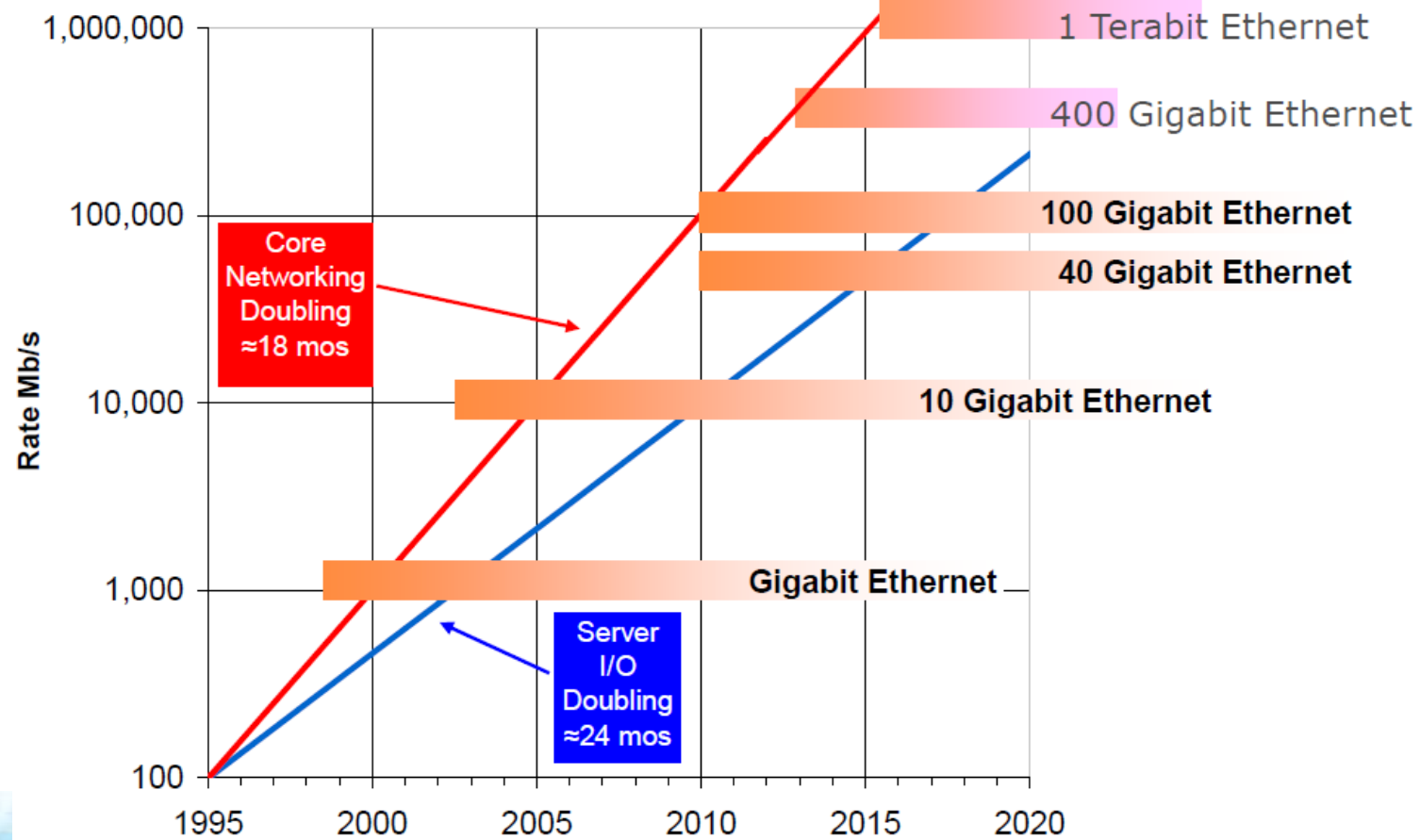


100G konečně realitou

Co a proč měřit na úrovni 100G

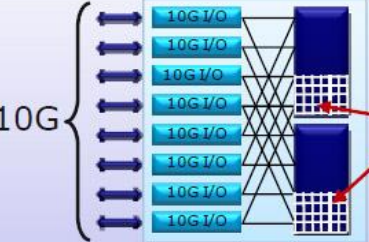
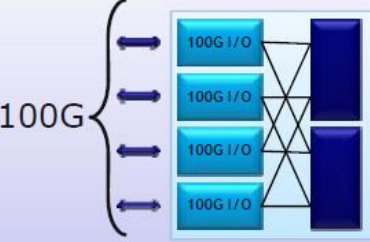
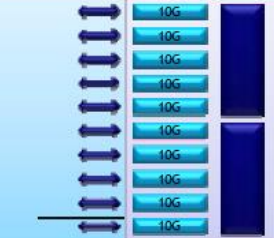
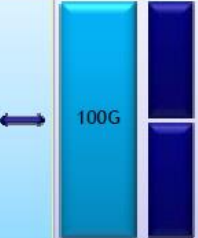
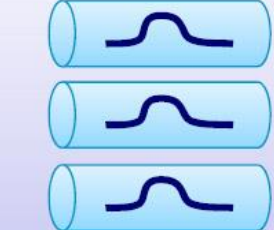
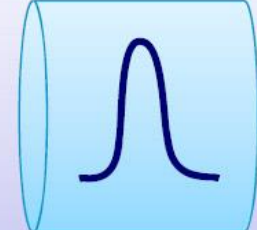
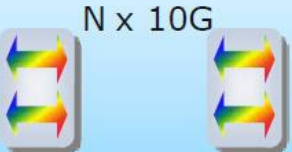

Nárůst objemu přenášených dat



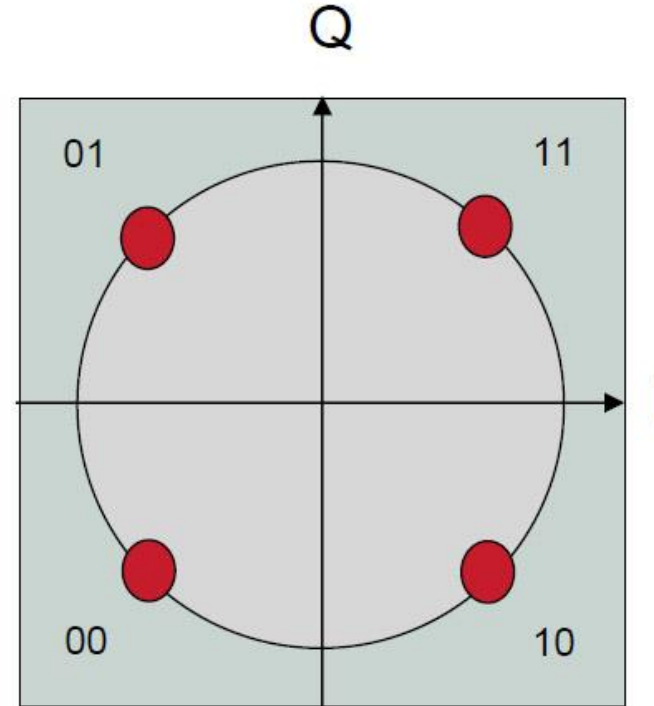
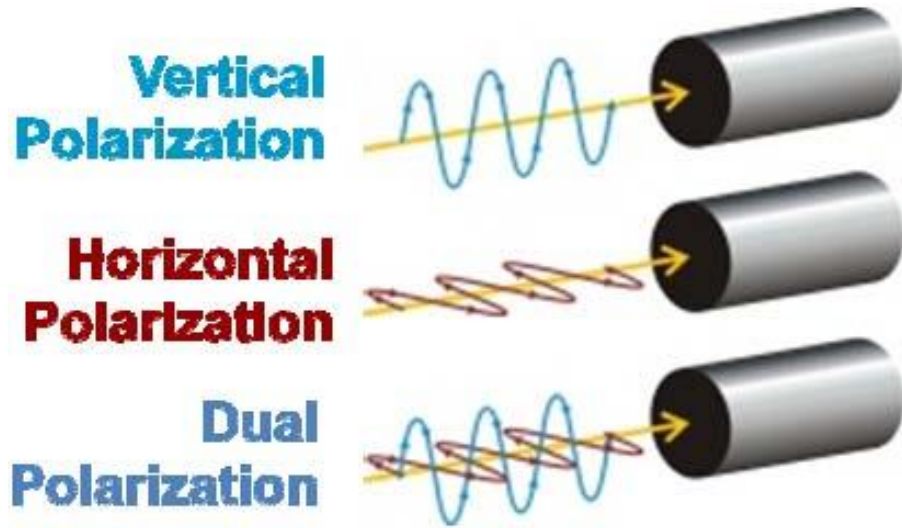
Jak jsme dosud zvyšovali kapacitu

- SDM – více vláken, stejná rychlost (ale vyšší celkové náklady na instalaci a správu vláken + potřeba optických zesilovačů pro delší úseky)
- TDM – vyšší rychlost na stejném počtu vláken ale problémy s PMD & CD, náročná konstrukce budičů laserů
- WDM – stejná rychlost i stejný počet vláken (ale nároky na zesilovače s vyrovnanou přenosovou charakteristikou, drahé lasery)

10x 10G vers. 1x 100G

10G interfaces	100G interfaces
 <p>10G {</p> <p>On high-end core routers - Fabric capacity is poorly utilized with lower speed I/O.</p>	 <p>100G {</p> <p>Better utilization of fabric bandwidth due to improved access into the fabric.</p>
 <p>On high-end core routers: 10 x 10G interfaces subscribe 100G of bandwidth.</p>	 <p>The price of 100G interface on router will be much lower than 10 x 10G for the same capacity.</p>
 <p>Small capacity wavelengths limit bandwidth bursts or increases the need for Link Aggregation Groups.</p>	 <p>A high capacity wavelength is better able to accommodate peak bandwidth limiting the need for Link Aggregation Groups.</p>
 <p>N x 10G</p> <p>Large wavelength count consumes the DWDM grid resulting in a higher network complexity and cost.</p>	 <p>100G</p> <p>Efficient use of DWDM grid. No parallel wavelengths means lower OAM cost and fewer managed entities.</p>

Redukce symbolové rychlosti



Dosud používané kódování NRZ / RZ nelze použít

Oddělené využití polarizačních složek

Vícestavová fázová modulace -> snížení symbolové rychlosti 4x, zlepšení poměru S/N, snížený vliv PMD/CD

DP-QPSK, koherentní přijímač

Dual Polarization QPSK (DP-QPSK) modulace navržená OIF je klíčovou technologií umožňující velké překlenutelné vzdálenosti na linkové straně

Dva ortogonální optické signály se shodnou frekvencí

Jediný vysílací laser, každá polarizovaná složka nese polovinu datového obsahu

Dvě polarizace = nižší řád modulace = menší optická šířka pásma

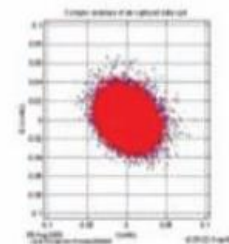
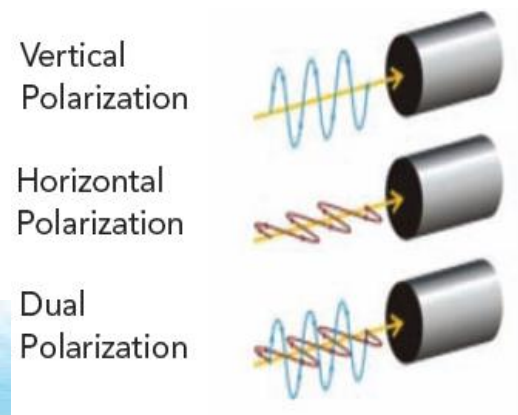
Přenáší 100G užitečného obsahu + záhlaví v 50GHz spektru

Detekce na koherentním přijímači

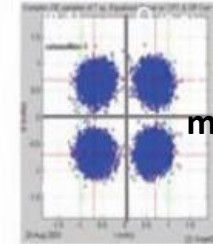
Frekvence místního oscilátoru přijímače je na stejné frekvenci jako přijímaný signál, Odděluje jednotlivé složky polarizace, detekuje bity QPSK

Zachovává informaci o amplitudě a fázi, umožňuje kompenzaci obou disperzí v rámci DSP zpracování

**Dvojitá
polarizace**



Rx Data
before DSP



Rx Data
after DSP

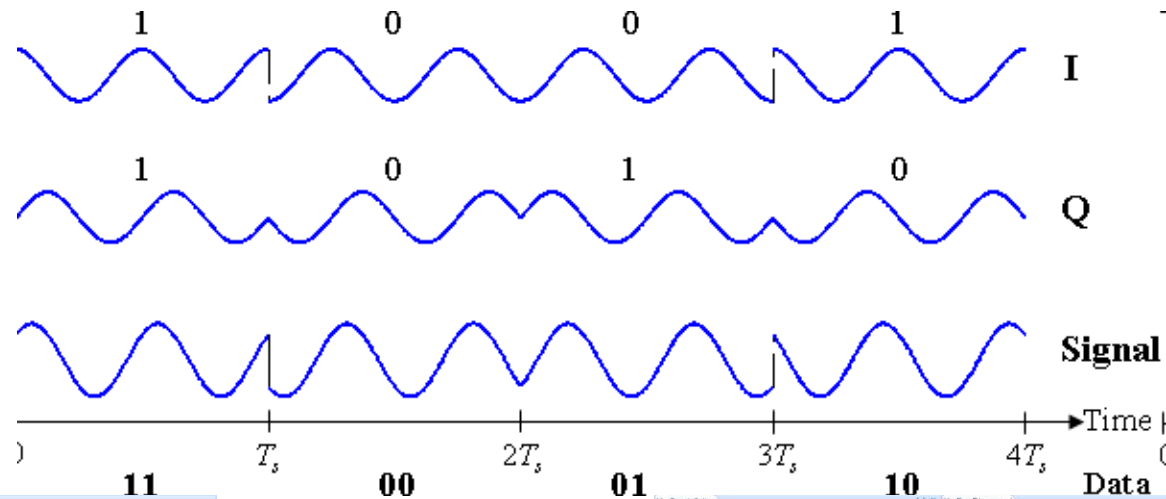
**Kvadrurní
modulace s fázovým
posuvem**

Koherentní přijímač

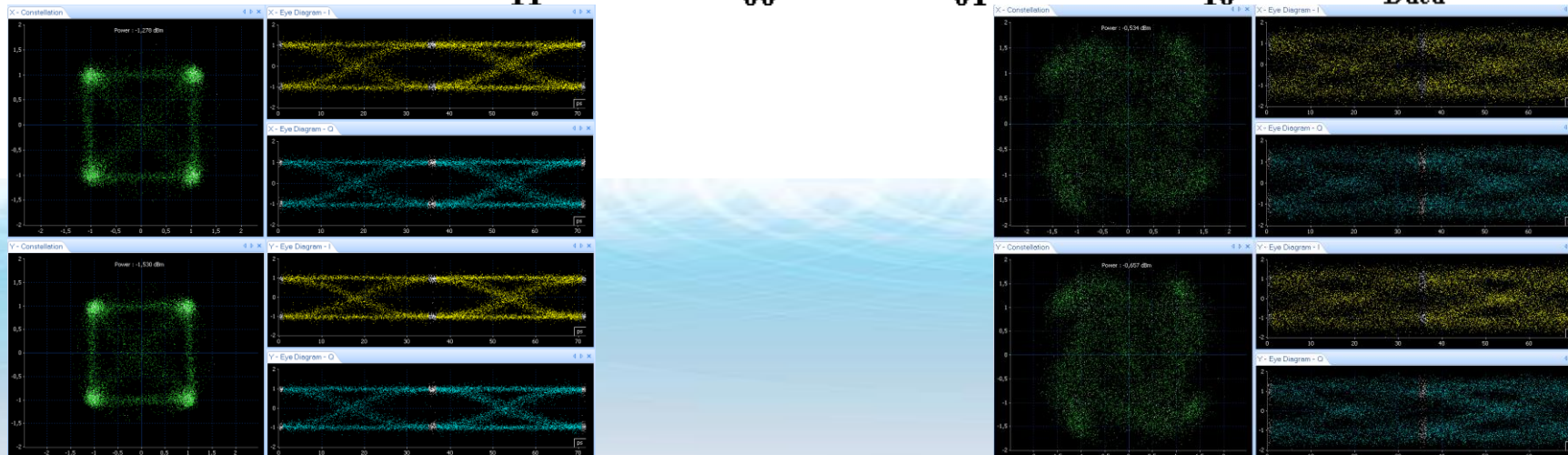
Informace je nesená fázovými změnami

Zlepšený poměr S/N , vyšší odolnost proti vlivu CD / PMD

A/D převodníky pro rychlost 28/56 Gbit/s



IEEE P802.3ba StandardP

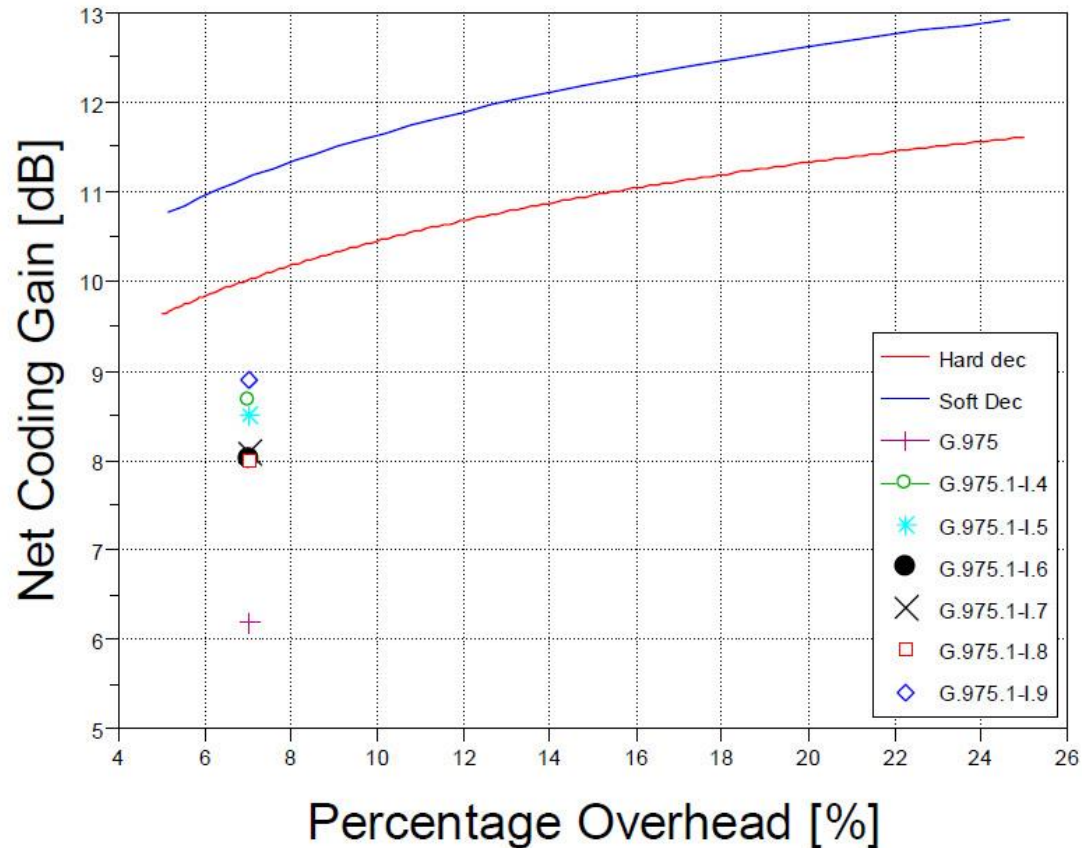


Korekce chyb v dopředném směru

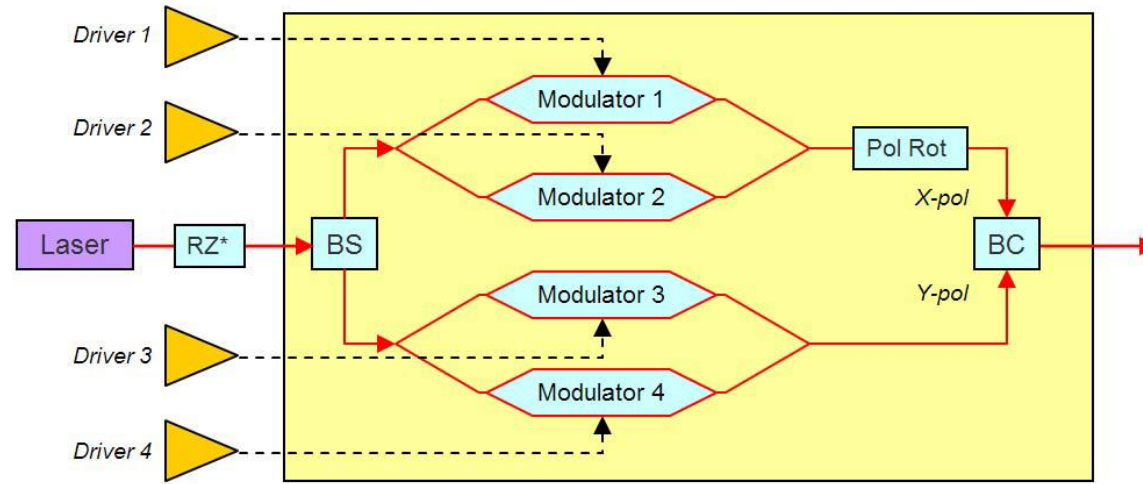
RS (255,239), účinnost jednotlivých verzí

Konstelační diagram

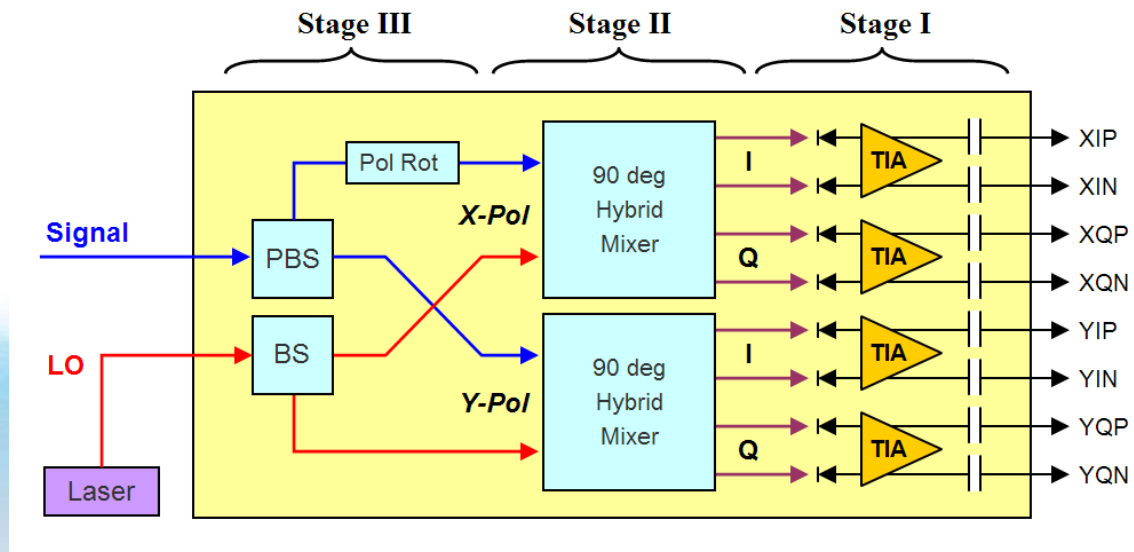
Diagram oka



Vysílač a detektor



* Optional RZ Carver



Sériově – paralelní převod

100GE Protocol Stack

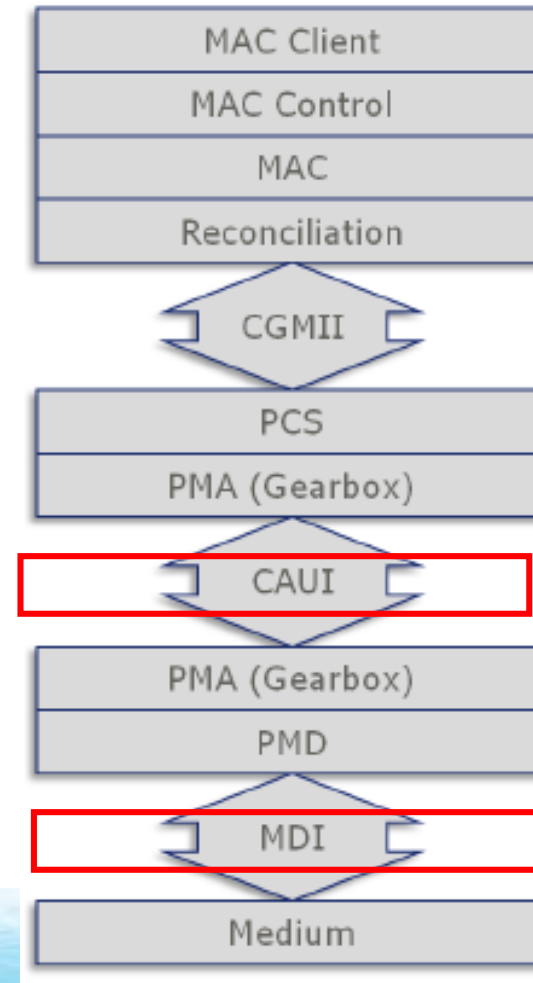
Jediná technologie před 100G využívající paralelní struktury byla linková agregace

100GE přidává do přenosového řetězce převodní blok na úrovni podvrstvy PMA. Je to mezivrstva mezi úrovněmi PCS a PMD

V anglické terminologii se vžila přezdívka “gearbox”, neboť hlavní funkcí této mezivrstvy je multiplexní převod mezi řetězcí PCS a PMD v obou směrech.

Tato multiplexní funkce zajišťuje možnost přenosu datového toku Ethernet v jednotlivých kanálech

Zatímco datová vrstva MAC běží na plné rychlosti 100Gbit/s, vlastní přenosové médium může využívat větší počet nezávislých paralelních datových toků bez ztrát a snížení účinnosti.



Physical Coding Sublayer

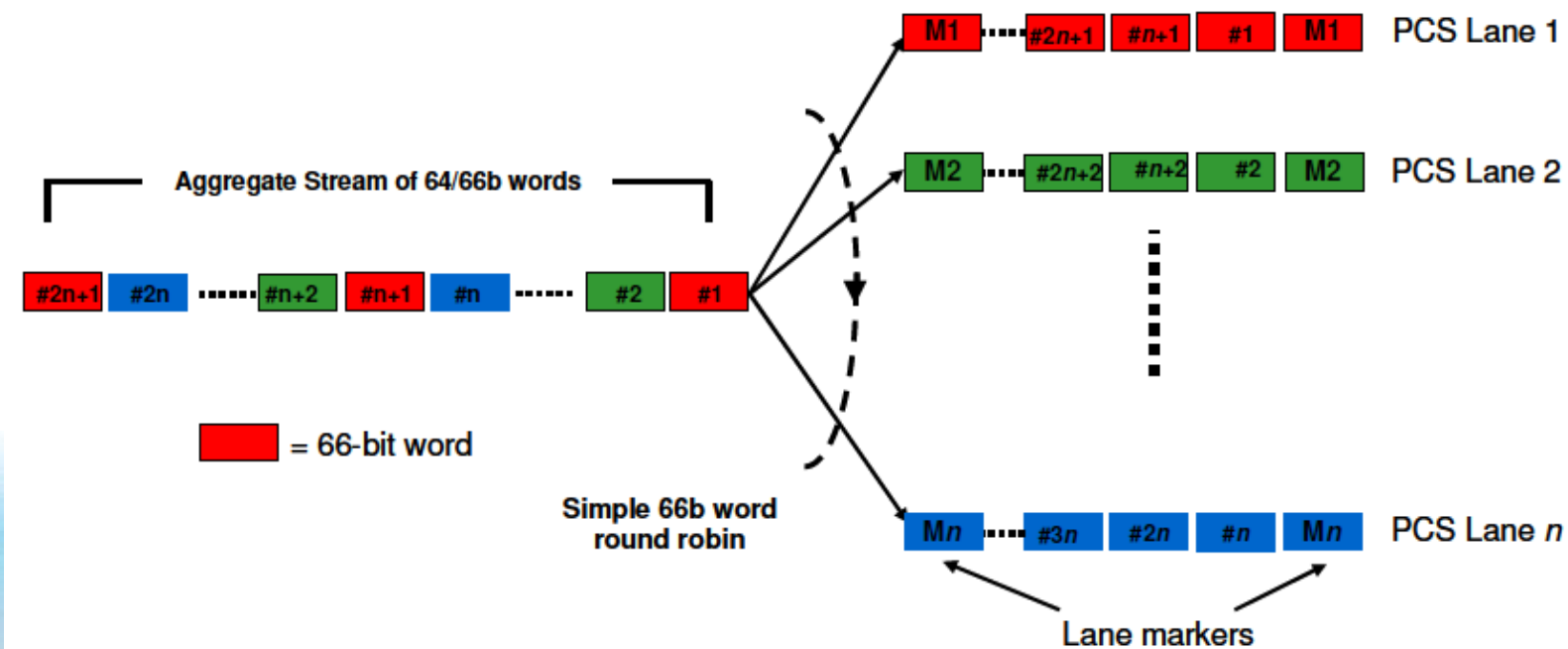
Podvrstva Physical Coding Sub-layer (PCS)

Převádí data z jednotlivých rozhraní Media Independent Interface (MII) na podvrstvu PMA.

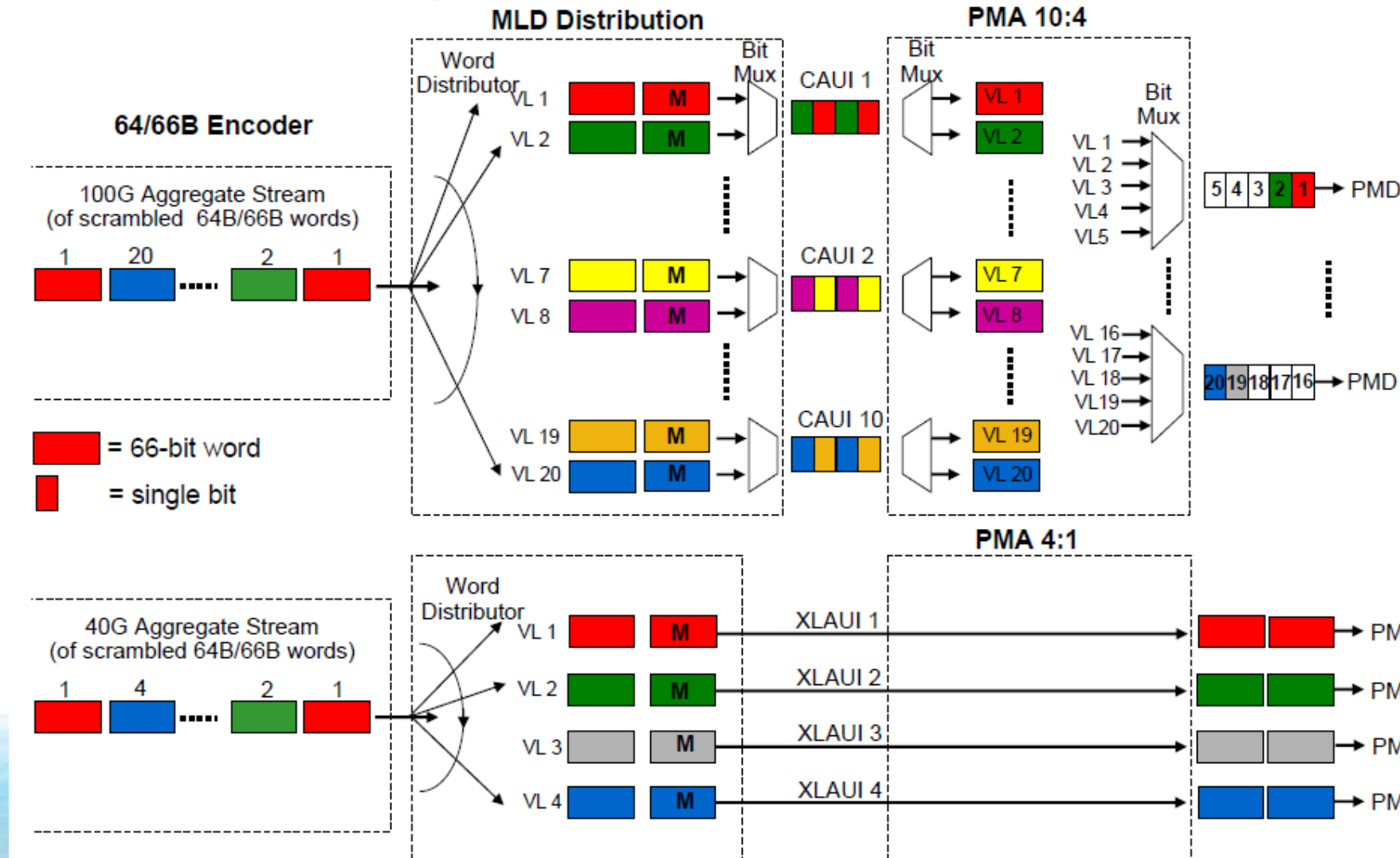
Agregovaný tok z MII je kódován do 64B/66B schématu (stejně jako 10GE)

Cyklická distribuce 66-bit bloků zajišťuje jejich rozdělení do většího počtu paralelních cest nazvaných "PCS Lanes," z nichž každá je vždy periodicky označena vlastním markerem

40GE využívá čtyři PCS cesty, 100GE pak dvacet PCS cest



Multiplexní schema

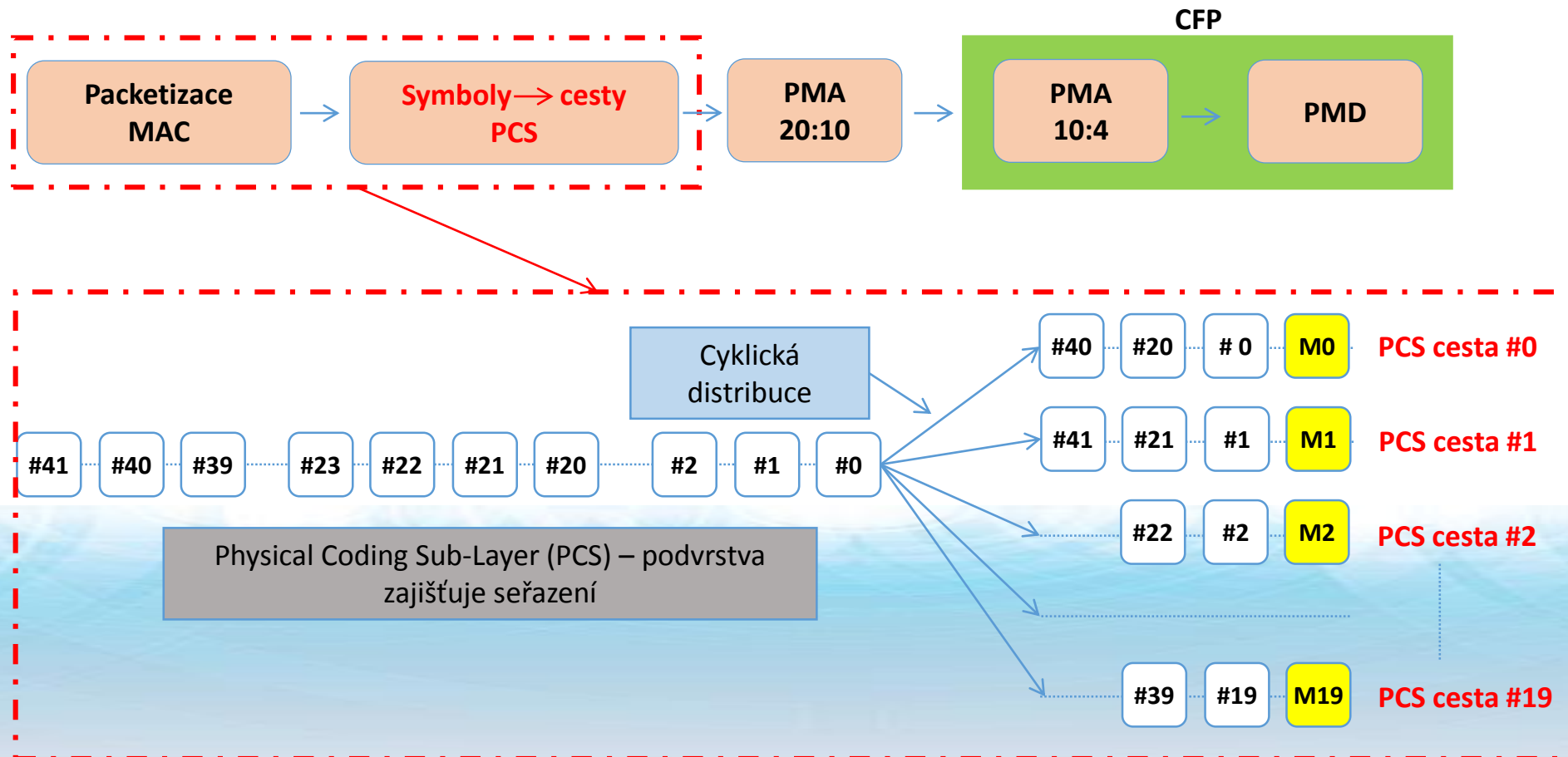


Koncept paralelních cest PCS

Přenos 100G využívá paralelních struktur na více úrovních

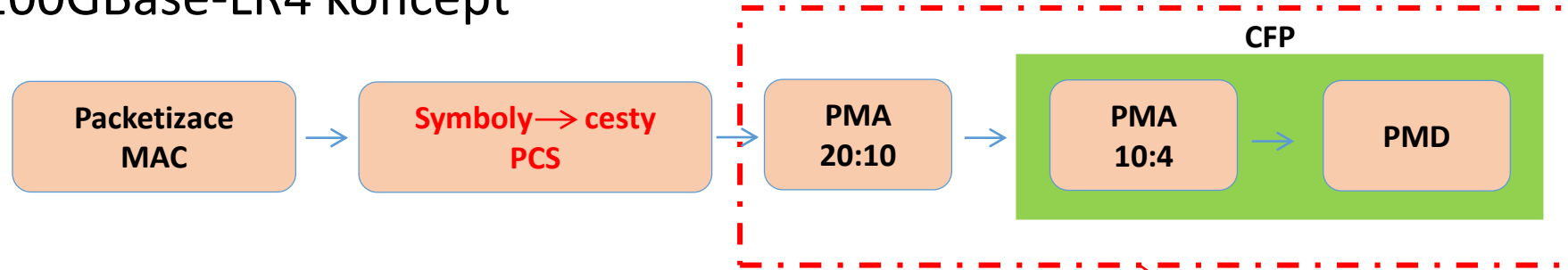
Krok #1: přidá se záhlaví MAC a seskupí se vždy skupiny po 8 bytech do 64/66B symbolů

Krok #2: symboly se cyklicky rozdělují do PCS #0 cest a přidají se značky M0 PCS cest každých 210us, které zajistí správné seřazení



Transcievery CFP, CFP2, QSFP+...

100GBase-LR4 koncept

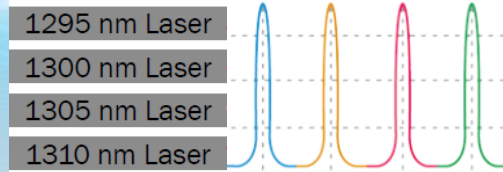
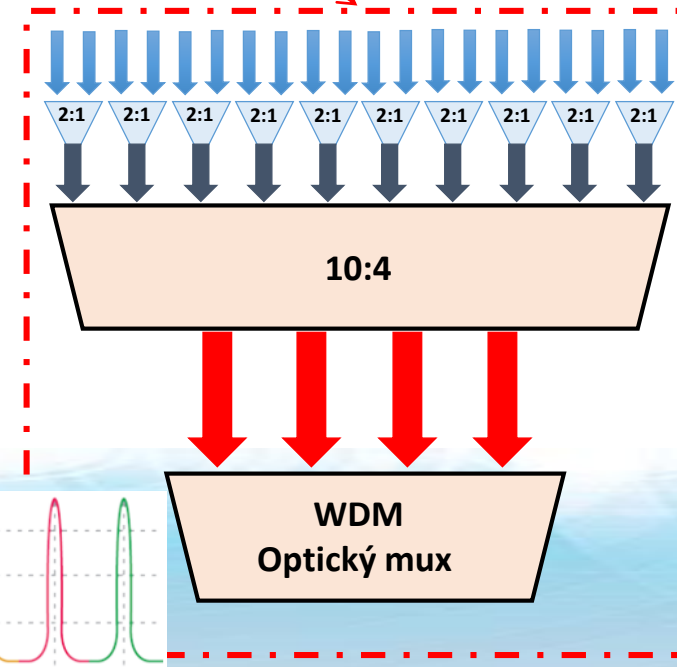


Krok #3: Multiplex 20 PCS cest do 10 CAUI cest

Krok #4: Multiplex 10 CAUI cest do 4 PMD cest

Krok #5: Převod 4 PMD cest na optické cesty s NRZ kódováním

Krok #6: Multiplex 4 optických cest do WDM signálu

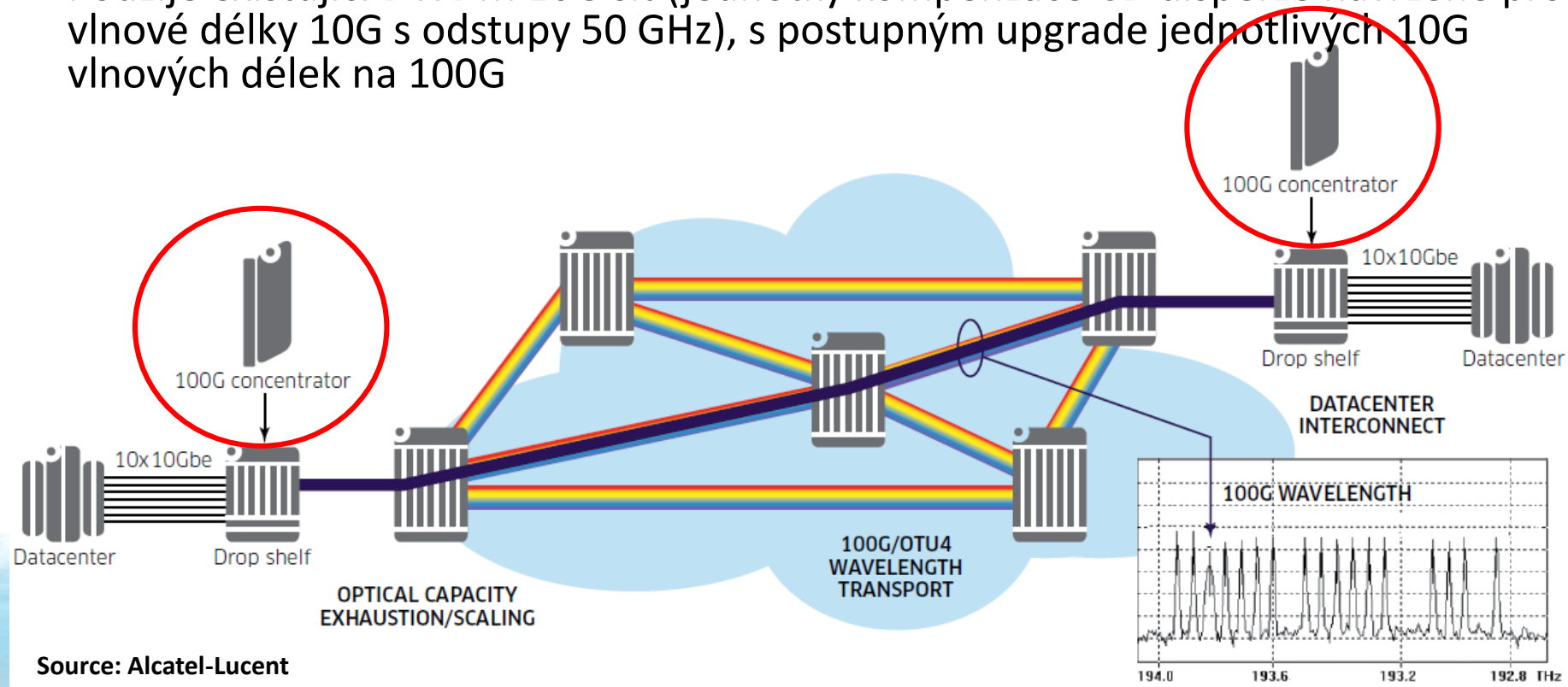


Využití fyzické vrstvy 10G

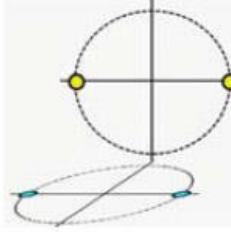
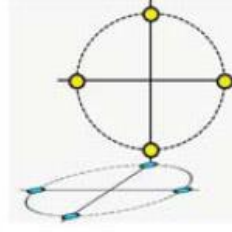
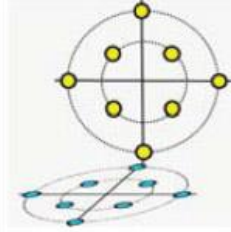
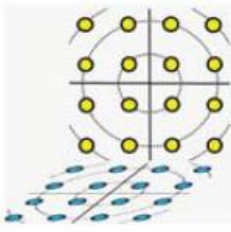
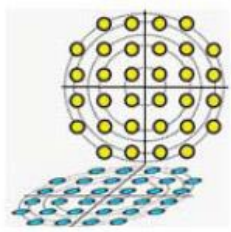
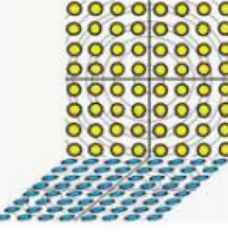
Úkol č.1: Propojení datových center s použitím 100G koncentrátorů a zvýšení celkové kapacity

100G zlepšuje efektivitu nákladů (CAPEX/OPEX) v porovnání s 10x10G toky

Použije existující DWDM 10G síť (jednotky kompenzace CD disperze navržené pro vlnové délky 10G s odstupy 50 GHz), s postupným upgrade jednotlivých 10G vlnových délek na 100G



Vyšší modulační schemata

Modulation format	PM-BPSK	PM-QPSK	PM-8QAM	PM-16QAM	PM-32QAM	PM-64QAM
bits/Symbol	2x1	2x2	2x3	2x4	2x5	2x6
Constellation						
OSNR penalty (dB)	0	0	2	4	6	8.5

Modulation format	PM-QPSK	PM-16QAM	PM-QPSK	PM-8QAM	PM-16QAM	PM-32QAM	PM-64QAM	PM-256QAM	PM-32QAM	PM-64QAM
Bitrate (Gb/s)	100	200	400					1000		
Symbol Rate (Gbd)	28-32	28-32	112-128	75-85	56-64	45-51	37-43	28-32	112-128	93-107
Bits/Symbol	4	8	4	6	8	10	12	16	10	12
Channel Spacing ¹ (GHz)	50	50	200	133	100	80	67	50	200	166
SE ¹ (bits/s/Hz)	2	4	2	3	4	5	6	8	5	6
no. of C-band channels	88	44	22	33	44	55	66	88	22	26
Total Capacity (Tb/s)	8.8	17.6	8.8	13.3	17.6	22	26.4	35	22	26
OSNR ² (dB) @ min Baudrate	12.2	19.2	18.2	20.2	22.2	24.2	26.7	> 30	28.2	30.7
OSNR ² (dB) @ max Baudrate	9.8	16.8	15.8	17.8	19.8	21.8	24.3	> 30	25.8	28.3
Penalty vs. 100G (dB)	0	7	6	8	10	12	14.5	> 20	16	18.5
1) same margin to filtering as 100G (f _n =1.56)										
2) ref to theoretical 40 Gb/s values										

Měřicí úlohy na 100G

- Na rozdíl od všech předchozích technologií je důležité měřit na Layer 1 – jde o konsistenci signálu. Konstelační diagram, Eye diagram, kontrola disperzí ve fázi výstavby a uvedení do provozu.
- První zátěžové testy – Line rate BERT, Lane Alignment & Skew Management
- Druhý krok: management alarmů, simulace rozpojení (APS), test multiplexních schemat na vrstvě OTU – ODU0 / ODUflex do OTU3 a OTU4
- Třetí krok: Ethernet Transmission Performance Analysis – standardní testy Y.1564, RFC2544, Traffic Loading, Frame Loss, Packet Latency, Packet BER, Throughput, Errors/Alarms, Signal Capture, Signal Monitoring *pro celou škálu velikosti rámců.*
- Testy po dokončení instalace v celé síti: QoS a Vperf na 10G / 1G s přenosem přes 40G/100G

100/40 Gbps Ethernet - Are you ready?

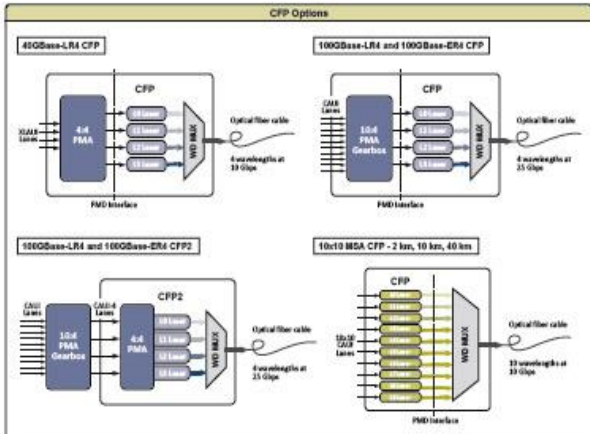


PMD LAYER

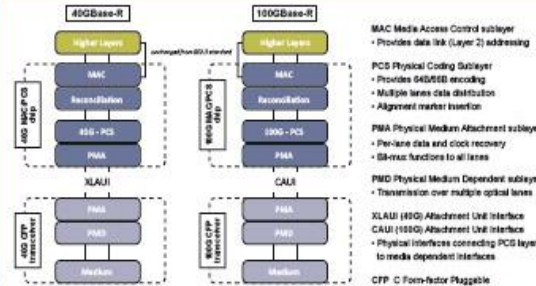
The Physical Medium Dependent sublayer or PMD defines the details of transmission and reception of individual bits on the physical medium. The PMD layer interfaces with the PMA layer and translates the encoded data to and from signals suitable for the physical medium. The PMA continuously sends parallel bit streams to the PMD, one per lane. The PMD converts these streams of bits into separate optical signal streams; the optical signal streams are then wavelength division multiplexed and delivered to the Medium Dependent Interface (MDI).

	CFP Types							
	40GBase-LR4	100GBase-DR10	100GBase-LR4	100GBase-ER4	10x10 - 2 km	10x10 - 10 km	10x10 - 40 km	
Standard	IEEE 802.3ba	IEEE 802.3ba	IEEE 802.3ba	IEEE 802.3ba	10x10 MSA	10x10 MSA	10x10 MSA	
Wavelength	1310 nm WDM (ITU-T G.694.2)	850 nm (10 fiber pairs)	1310 nm WDM (ITU-T G.694.1)	1310 nm WDM (ITU-T G.694.1)	1550 nm WDM	1550 nm WDM	1550 nm WDM	DWDM (ITU-T G.694.1)
Reach	10 km	100 m	10 km	40 km	2 km	10 km	40 km	
Optical rate	4 x 10 Gbps	10 x 10 Gbps	4 x 25 Gbps	4 x 25 Gbps	10 x 10 Gbps	10 x 10 Gbps	10 x 10 Gbps	

Lane	WDM Lane Assignments											
	Center Wavelength (nm)	Center Wavelength (nm)	Center Frequency (THz)									
	40GBase-LR4	100GBase-LR4/100GBase-DR4	10x10 MSA - 2 km	10x10 MSA - 10 km	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7	Band 8
L0	1271	1296.56	1523	196.10	196.00	195.90	195.80	195.70	195.60	195.50	195.45	195.40
L1	1291	1300.05	1531	196.00	196.00	196.00	196.00	196.00	196.00	196.00	196.00	196.00
L2	1311	1336.56	1539	195.90	195.80	195.70	195.60	195.50	195.45	195.40	195.35	195.30
L3	1331	1360.14	1547	195.80	195.70	195.60	195.50	195.45	195.40	195.35	195.30	195.25
L4	n/a	n/a	1555	195.70	195.60	195.50	195.45	195.40	195.35	195.30	195.25	195.20
L5	n/a	n/a	1563	195.60	195.50	195.45	195.40	195.35	195.30	195.25	195.20	195.15
L6	n/a	n/a	1571	195.50	195.40	195.30	195.25	195.20	195.15	195.10	195.05	195.00
L7	n/a	n/a	1579	195.40	195.30	195.20	195.15	195.10	195.05	195.00	194.95	194.90
L8	n/a	n/a	1587	195.30	195.20	195.10	195.05	195.00	194.95	194.90	194.85	194.80
L9	n/a	n/a	1595	195.20	195.10	195.00	194.95	194.90	194.85	194.80	194.75	194.70



IEEE 802.3ba STANDARD Protocol Stack



MAC Media Access Control sublayer
 • Provides data link (Layer 2) addressing

PCS Physical Coding Sublayer
 • Provides 64B/66B encoding
 • Multiple lanes data distribution
 • Alignment marker insertion

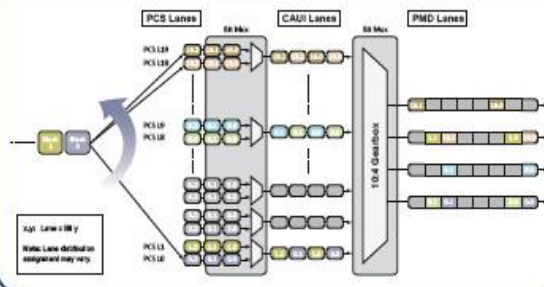
PMA Physical Medium Attachment sublayer
 • Per-lane data and clock recovery
 • Bit-error functions to all lanes

PMD Physical Medium Dependent sublayer
 • Transmission over multiple optical lanes

XAUI (80G) Attachment Unit Interface
CAUI (100G) Attachment Unit Interface
 • Physical interfaces connecting PCS layer to media dependent interfaces

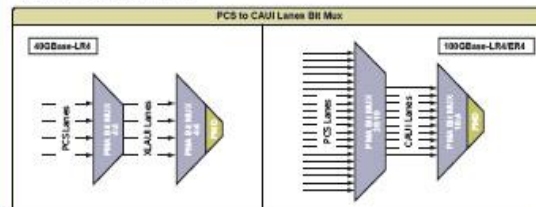
CFP: Form-factor Pluggable

THE BIG PICTURE 100GBase-R



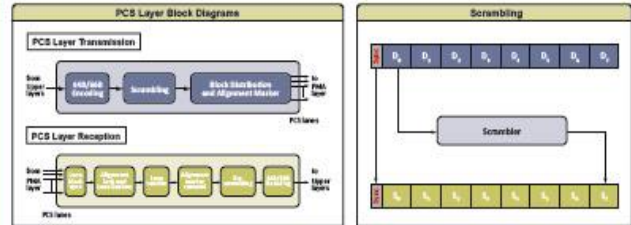
PMA LAYER

The Physical Medium Attachment sublayer or PMA allows the PCS to connect in a media-independent way with a range of physical media. It uses bit-level multiplexing to adapt the PCS lanes to the appropriate number of physical lanes required by the PMD interface.



PCS LAYER

The Physical Coding Sublayer (PCS) shields upper layers (MAC) from the specific nature of the underlying channel. Both 40GBase-R and 100GBase-R are based on a 64B/66B code which supports transmission of data and control characters, while maintaining robust error detection. When communicating with the lower layer (PMA), the PCS uses multiple serial streams or PCS lanes.



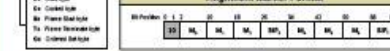
64B to 66B Blocks Encoding

Input Data	Block Payload	Output
8B Payload	0 1 2 ... 7	0 1 2 3 4 5 6 7
Control Word	8 9 10 11 12 13 14 15	8 9 10 11 12 13 14 15
64B Payload	16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64	16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64

Alignment Marker Encoding

8B Payload	Alignment Marker
0	0x00, 0x70, 0x40, 0x50, 0x60, 0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
1	0x10, 0x20, 0x30, 0x40, 0x50, 0x60, 0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
2	0x20, 0x30, 0x40, 0x50, 0x60, 0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
3	0x30, 0x40, 0x50, 0x60, 0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
4	0x40, 0x50, 0x60, 0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
5	0x50, 0x60, 0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
6	0x60, 0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
7	0x70, 0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
8	0x80, 0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
9	0x90, 0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
10	0xA0, 0xB0, 0xC0, 0xD0, 0xE0, 0xF0
11	0xB0, 0xC0, 0xD0, 0xE0, 0xF0
12	0xC0, 0xD0, 0xE0, 0xF0
13	0xD0, 0xE0, 0xF0
14	0xE0, 0xF0
15	0xF0

Alignment Marker Format



Multi-lane Distribution



PCS Alarms and Errors

Alarm/Event	Description
Loss of Block Lock (L0B)	Set after 80 invalid sync headers (80 or 11) in a 1024 window. Reset after 64 consecutive valid 68B sync headers are received (61 or 10).
Loss of Alignment Marker Lock (L0AM)	Set after 4 consecutive marker values are received that do not match the alignment marker that the lane is currently using to. Reset after 2 valid alignment markers are received 100B tracts apart.
Loss of Alignment (L0A)	Set if any of the PCS lanes are out of alignment marker lock or the incoming lanes skew exceeds the carrier compensation tolerance.
Invalid Sync Header (SH)	Set if an invalid sync header (80 or 11) is received.
Invalid Alignment Marker (AM)	Set if the received alignment marker does not match one of the alignment marker standard encoding type including 68B and 89B bytes.
BP	PCS lane bit-error-rate parity error.

