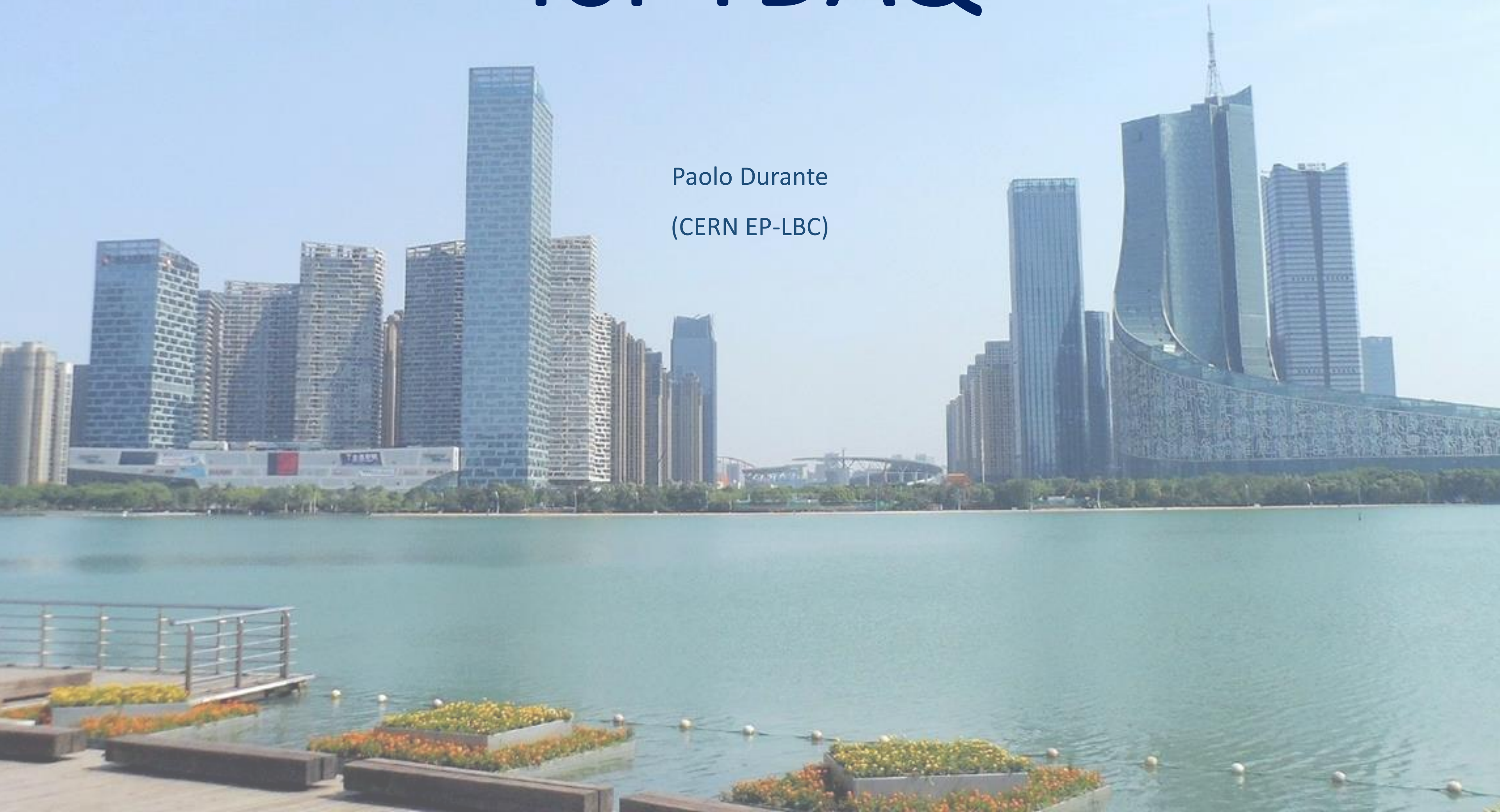
Optical Links for TDAQ

Paolo Durante

(CERN EP-LBC)



Outline

- Fundamental concepts
- Physical devices
- Applications
  - Frontend optical links
  - Backend optical links
- Installation / Commissioning

Why optical fibers (vs. copper)

PROs

- High rates (Tb/s)
- Long range (1000s Km)
- Low weight
- Low material budget
- No interference
- Improving technology
- Cheap material (silica)

CONs

- Expensive technology
- System complexity
- Mechanical fragility
- Installation complexity
- Termination complexity

Laser safety (1/2)

- Class 1
  - Safe
- Class 1M
  - Safe if not magnified
- Class 2
  - Visible light, triggers blink reflex safe ≤0.25s (e.g. laser pointers)
- Class 2M
  - Same as 2, if not magnified
- Class 3R
  - Low risk of injury, limit direct exposure (**typical tx**)
- Class 3B
  - Do not look directly (CD, DVD)
- Class 4
  - Permanent eye damage, skin burns, may light flammable materials

Laser safety (2/2)

Check your laser pointers before travelling...

Vincenzo Vagnoni <vincenzo.vagnoni@cern.ch>

Fr, 17.11.2023 15:17

An:lhcb-community (Group which contains all the actors of LHCb collaboration) <lhcb-community@cern.ch>

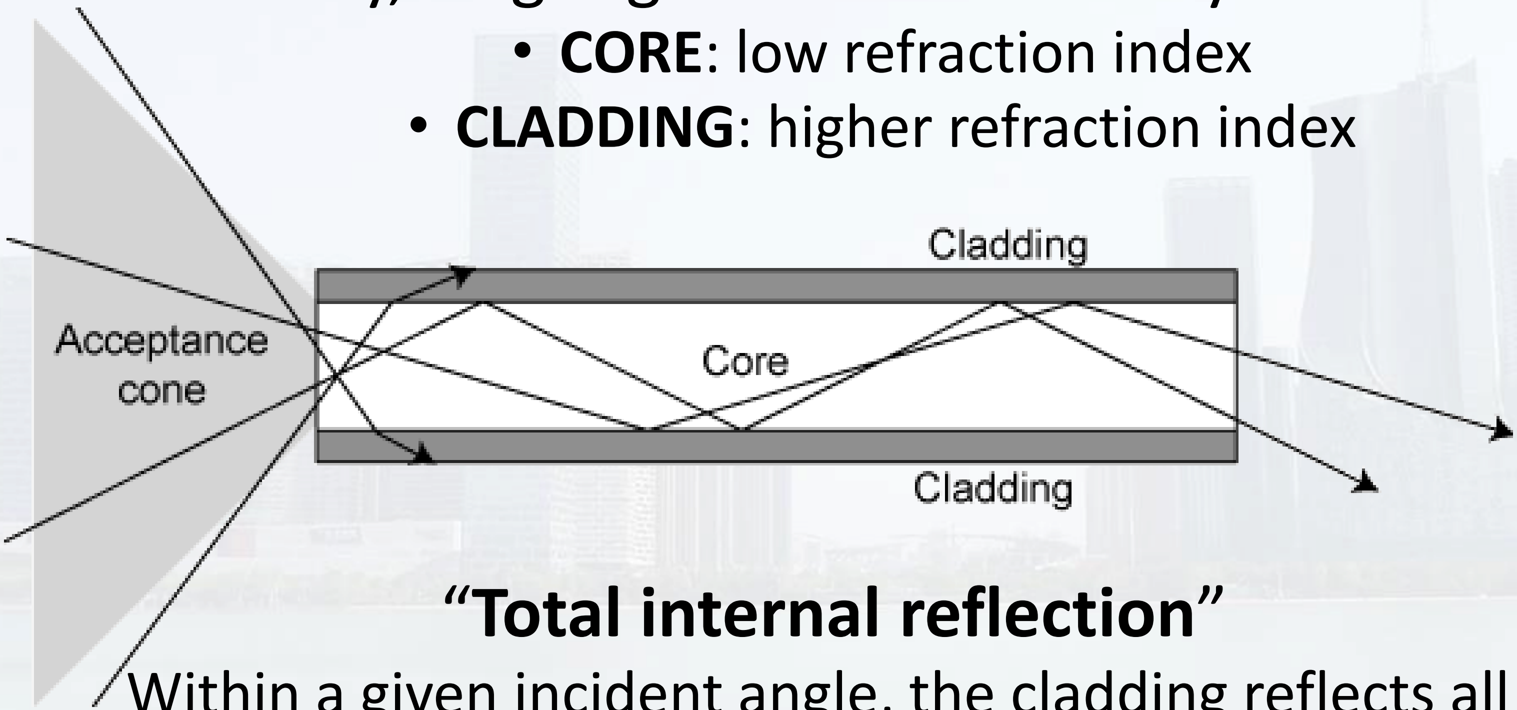
Dear colleagues,
we have been informed by the CERN HSE unit that there have been now 3 cases of people being arrested at the airport because they were carrying laser pointers with a forbidden power class. When travelling to/from CERN (and if you want also elsewhere), please check that, if you are carrying a laser pointer with you, its class is compliant with regulations.
Cheers,
Vincenzo

==
Vincenzo Vagnoni
—
Spokesperson LHCb experiment

How optical fibers work

Essentially, a light guide with a two layer structure

- **CORE**: low refraction index
- **CLADDING**: higher refraction index

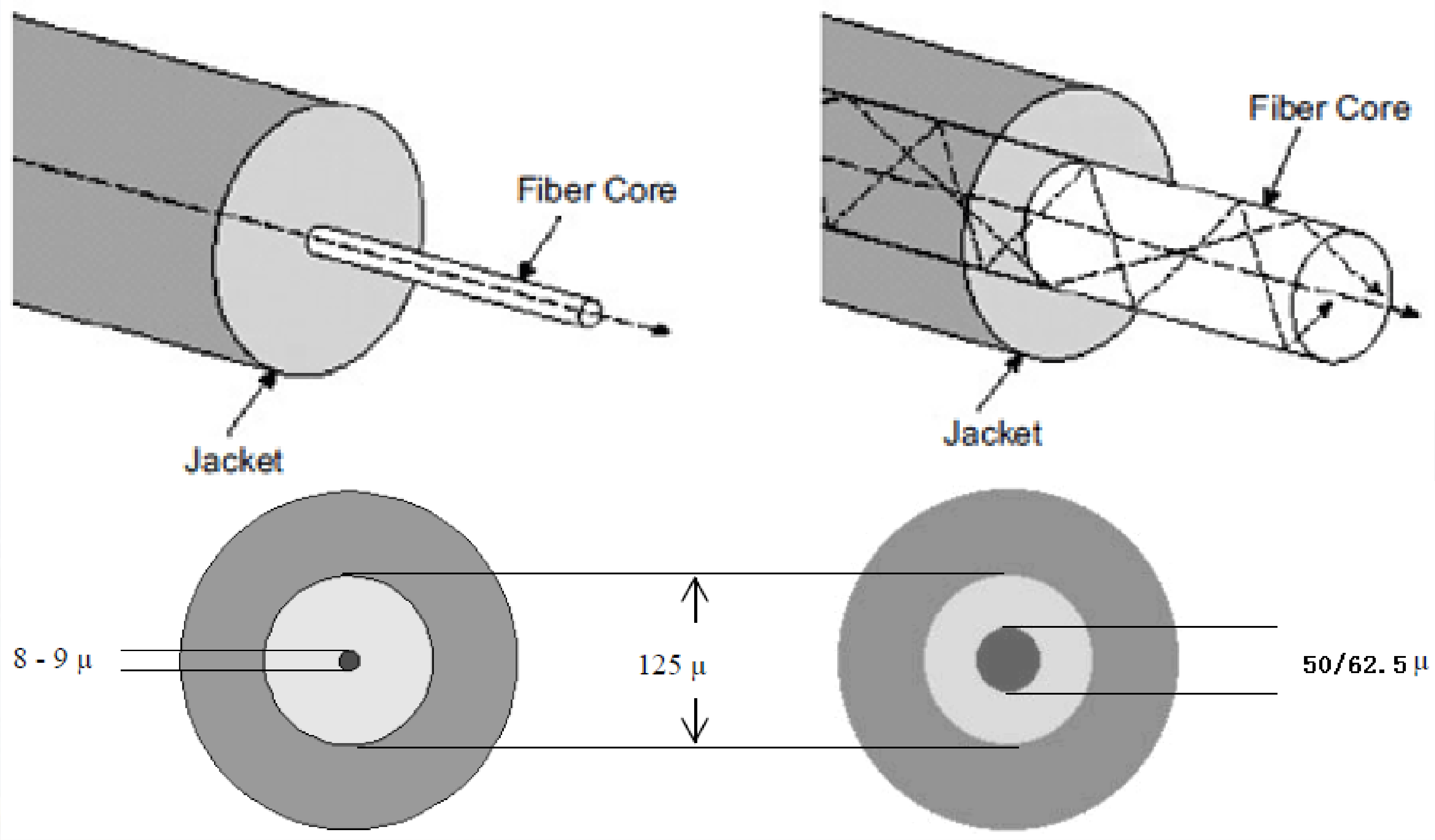


“**Total internal reflection**”

Within a given incident angle, the cladding reflects all light back into the core

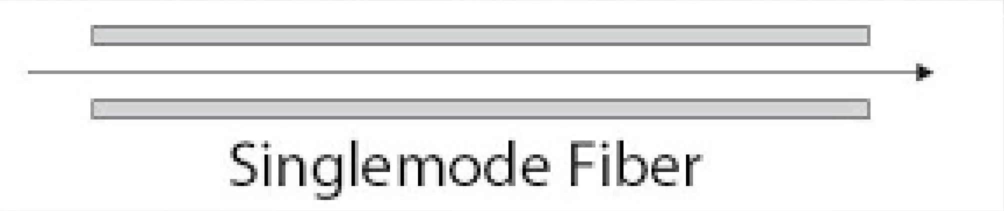
“Single-Mode” vs. “Multi-Mode”

Single-Mode Fiber (SMF) **Multi-Mode Fiber (MMF)**



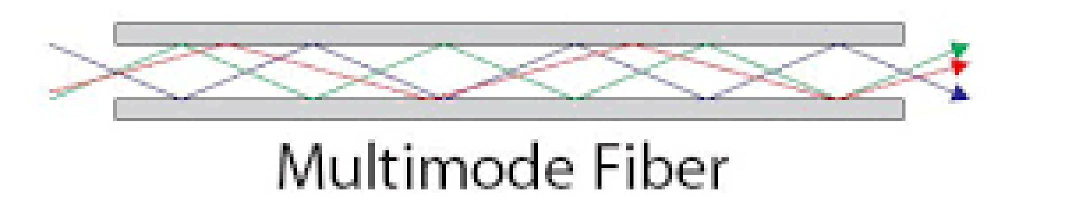
Single-Mode fiber (SMF)

- Smaller core diameter (8-10 µm)
- Higher bandwidth and longer range
- No “modal dispersion” (frequently used for WDM)
  - ~80 Km unamplified
  - 1000s of Km amplified
- Very expensive and precise laser transmitters

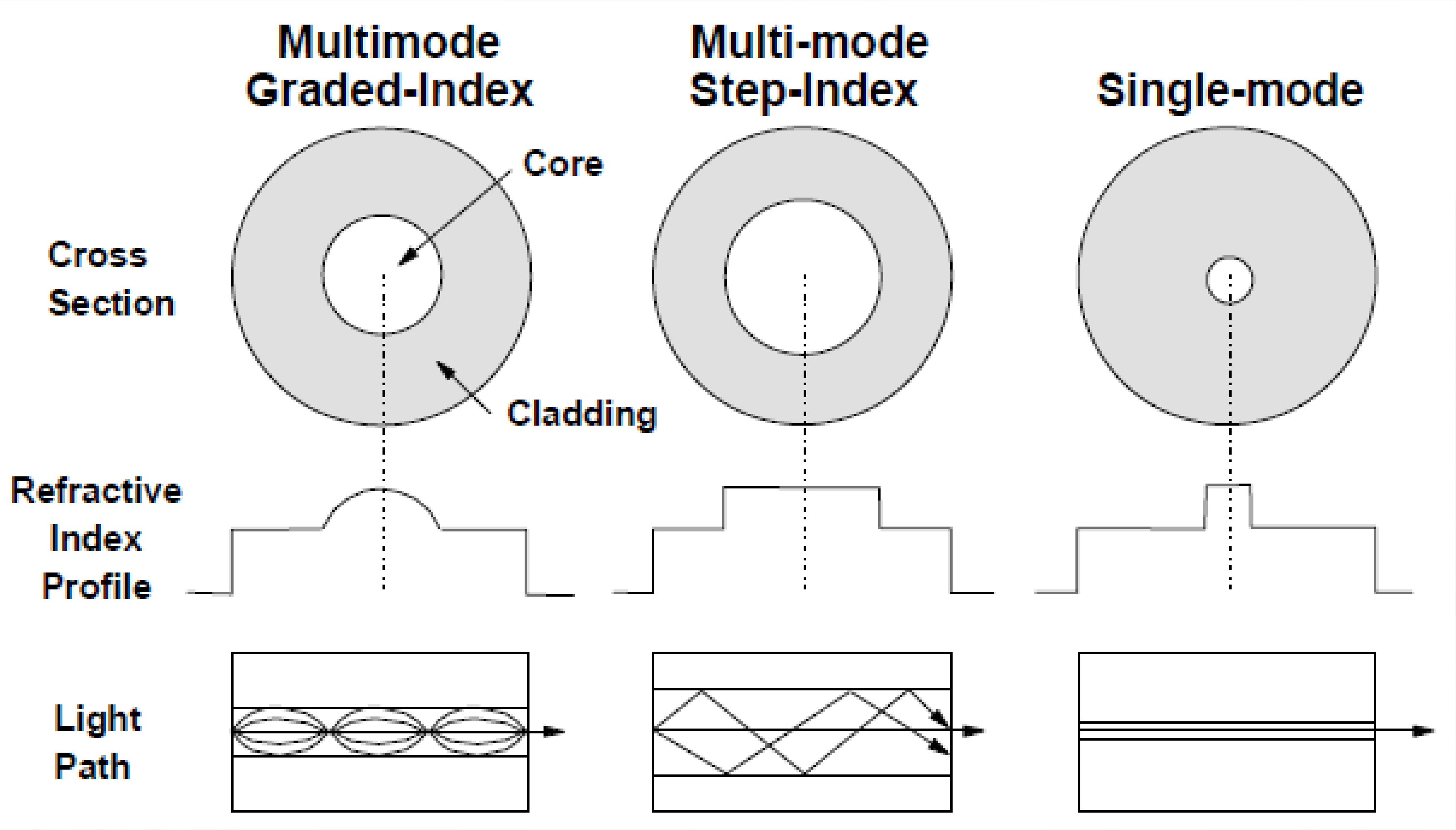


Multi-Mode Fiber (MMF)

- Much wider core (~50 µm)
- Much more forgiving wrt light quality & alignment
- Attenuation influenced by “modal dispersion”
  - Multiple propagation modes allowed in the light guide
  - Range limited to hundreds of meters
  - Less on graded-index wrt step-index fibers
- Light sources are considerably cheaper



Graded-index vs. Step-index fiber



OM3 vs. OM4

OM3

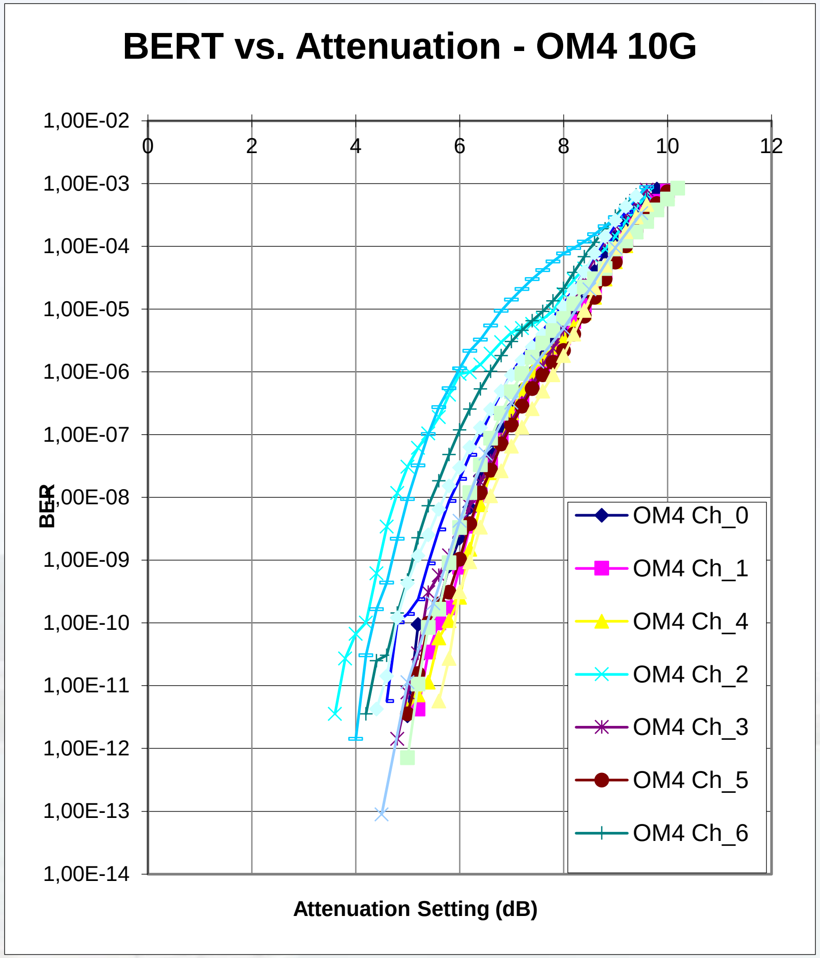
- <3.5 dB/Km @850nm
- EMB: 2000 MHz x km
- 1000m @ 1 Gb/s
- 300m @ 10 Gb/s
- 100m @ 100 Gb/s

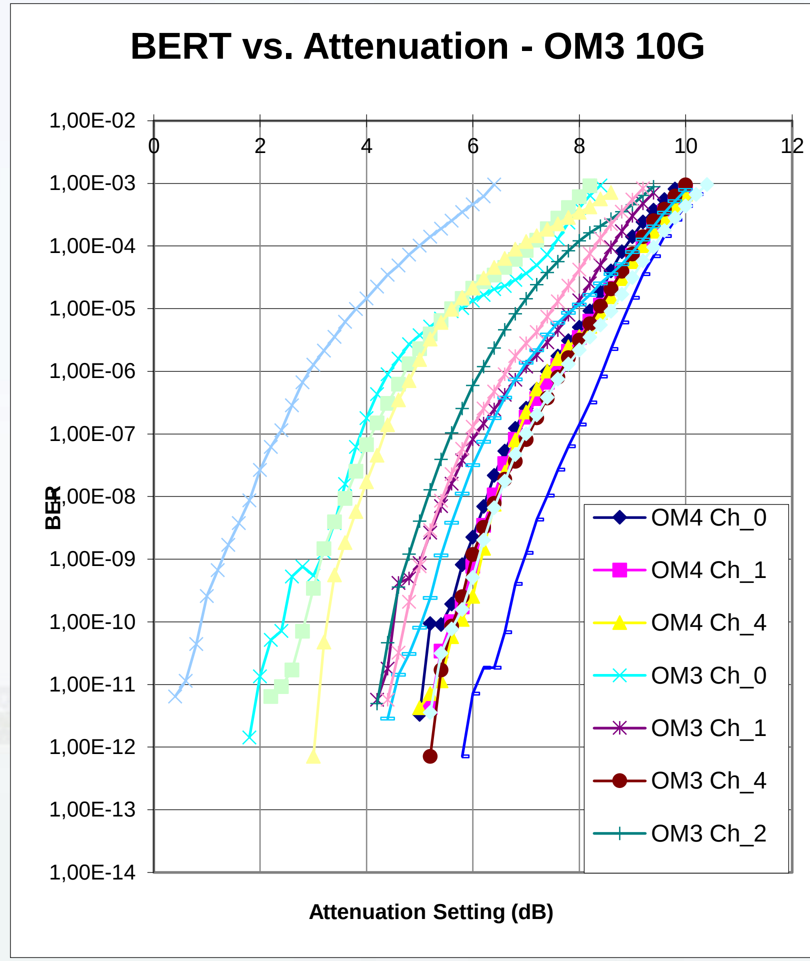
OM4

- <3.0 dB/Km @850nm
- EMB: 4700 MHz x km
- 1000m @ 1 Gb/s
- 500m @ 10 Gb/s
- 150m @ 100 Gb/s
- Higher cost

EMB = **Effective Modal Bandwidth** = max data rate at a given distance (constant) and λ
only accounts for modal dispersion, <u>not BER</u>

OM3 vs. OM4 Bit Error Ratio (BER)



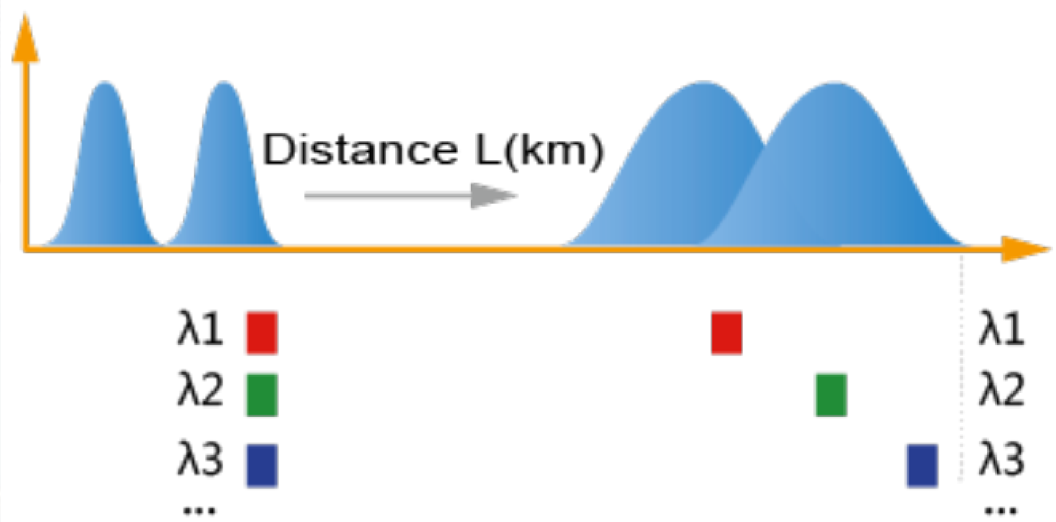


https://indico.cern.ch/event/297003

Ethernet: BER < 10^{-12}

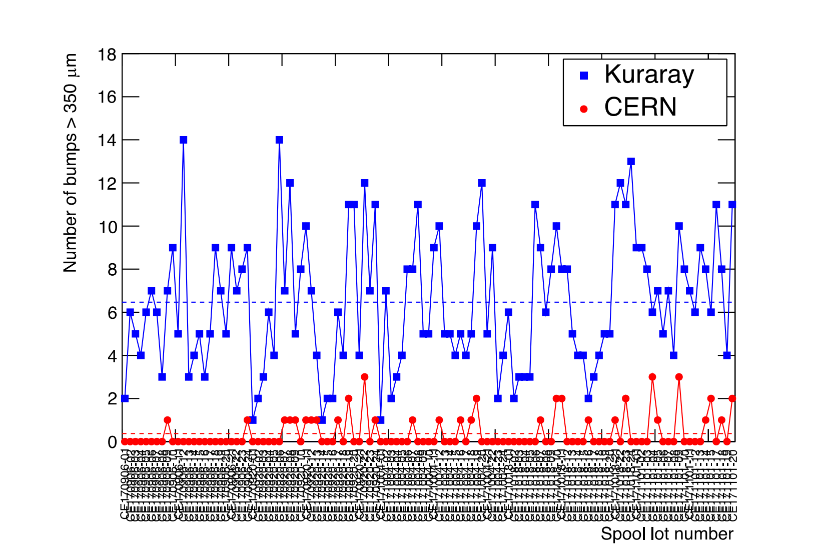
Chromatic Mode Dispersion (CMD)

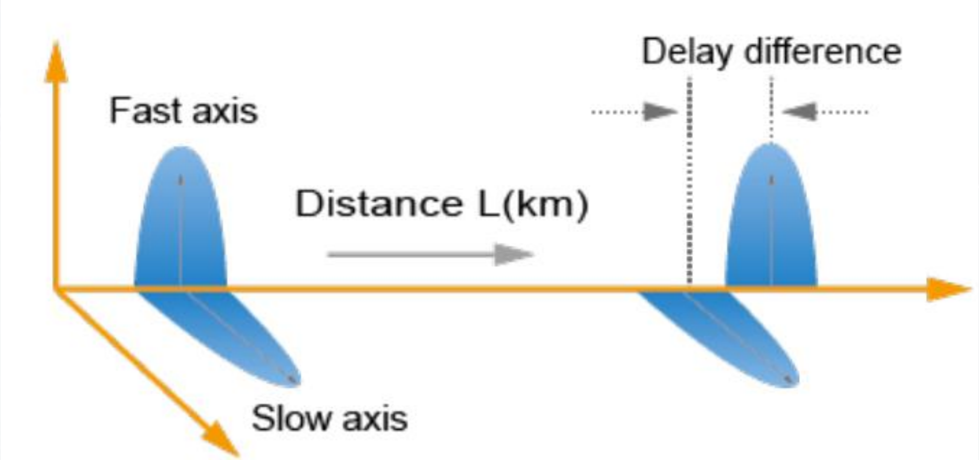
- Dispersion spreads out pulse shapes
- Different wavelengths (colors) propagate in a non vacuum at different speeds
- CMD is worse at higher data rates (f^2)



Polarization Mode Dispersion (PMD)

- Real optical fibers are not perfectly cylindrical
- Light with different polarization propagates at different speeds





Modulation

The simplest form of modulation is called “IM-DD”

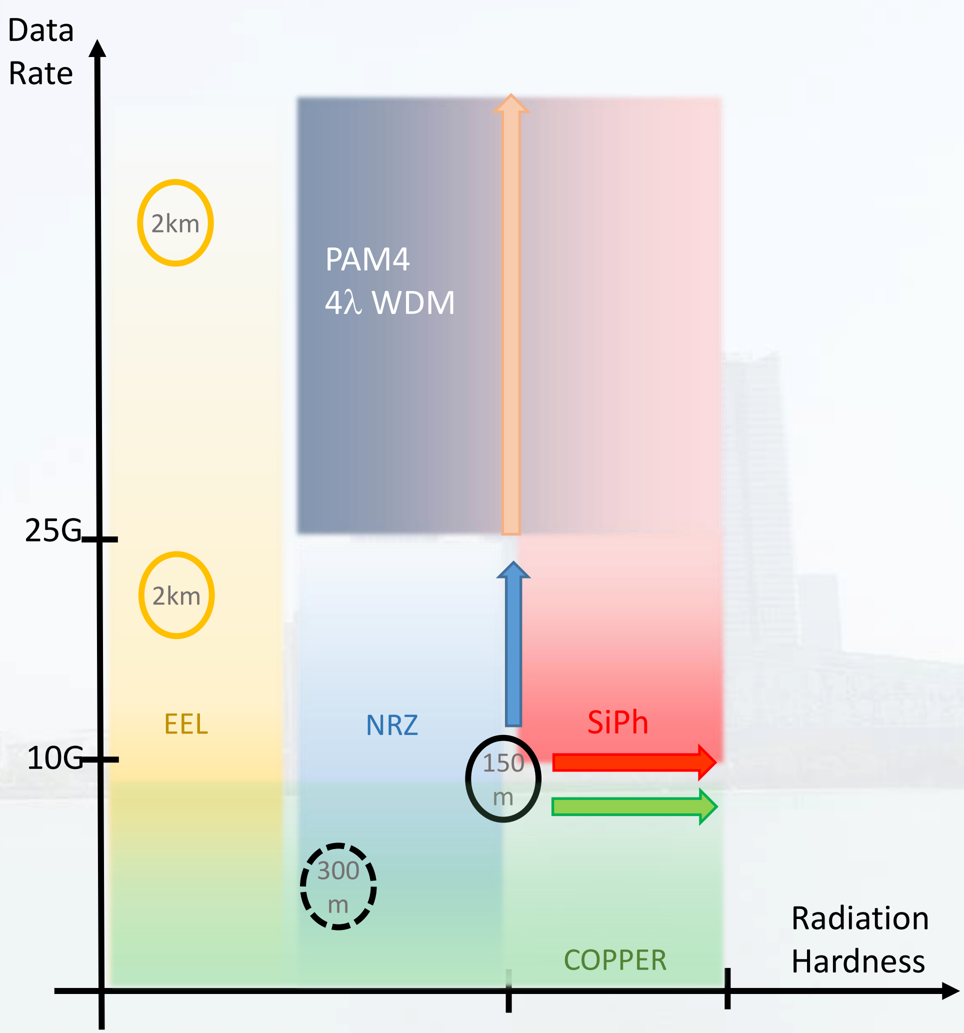
- Which stands for “Intensity Modulation with Direct Detection”
- Also called “Amplitude Shift Keying” (ASK)
- The most common version is “NRZ”, or “Non-Return to Zero”
- “Direct Detect” means only a photodiode is needed to RX

Historically, fiber optic systems were purely NRZ based

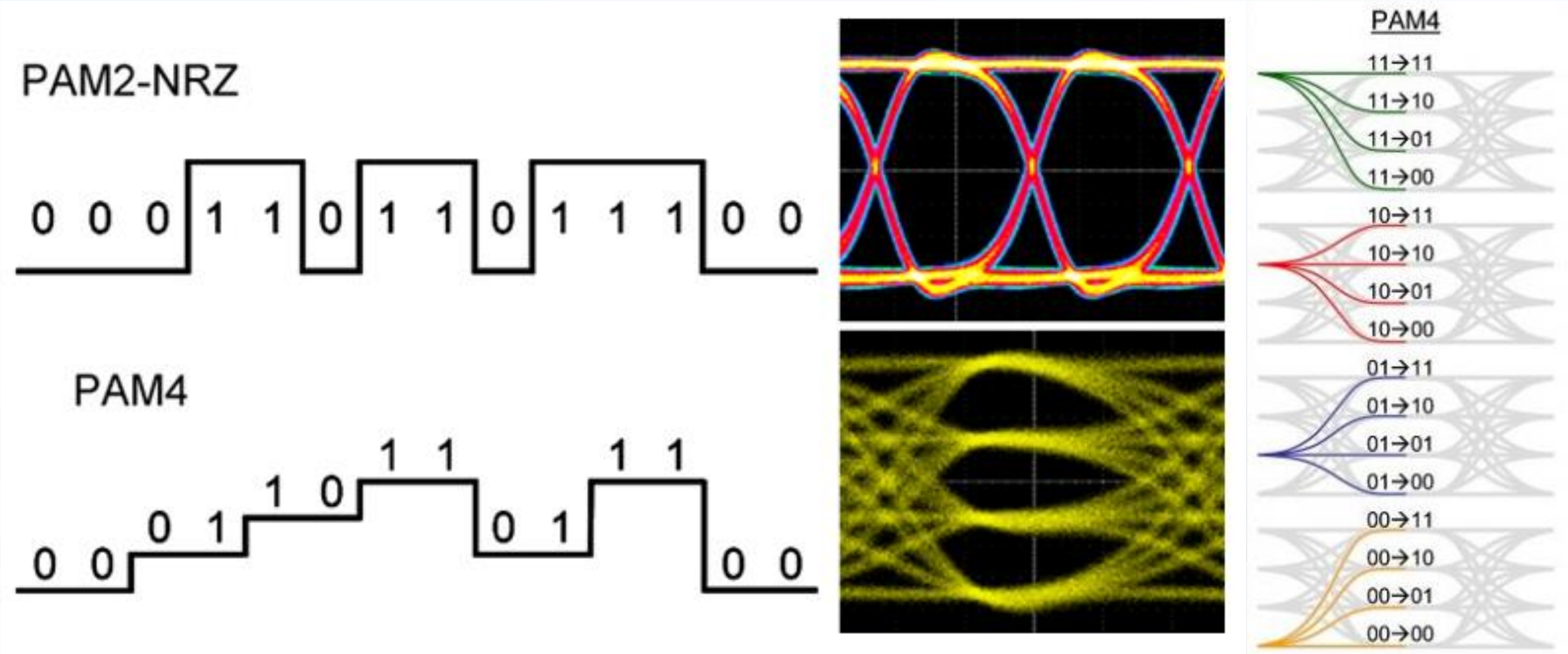
- All 10G and below optical technology is based around NRZ

How to go beyond NRZ?

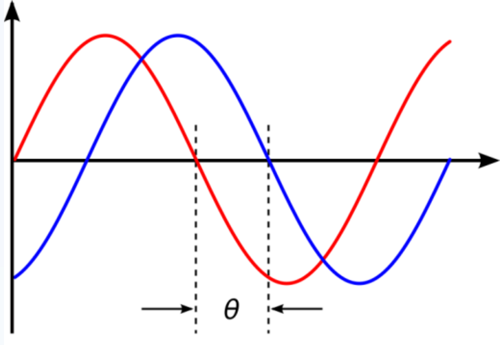
- More fibers in parallel
- DWDM over a single fiber
- Polarization Division Multiplexing (PDM)
- ASK with more amplitudes
- Modulate the “phase” of the signal over time (PSK)



Amplitude Shift Keying (ASK)



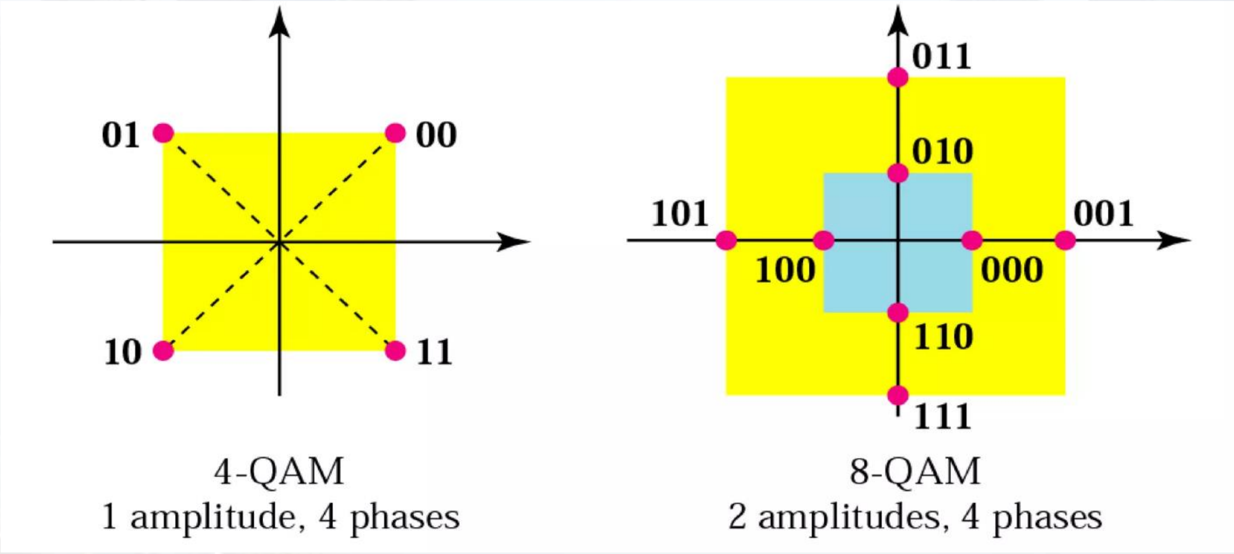
Phase Shift Keying (PSK)



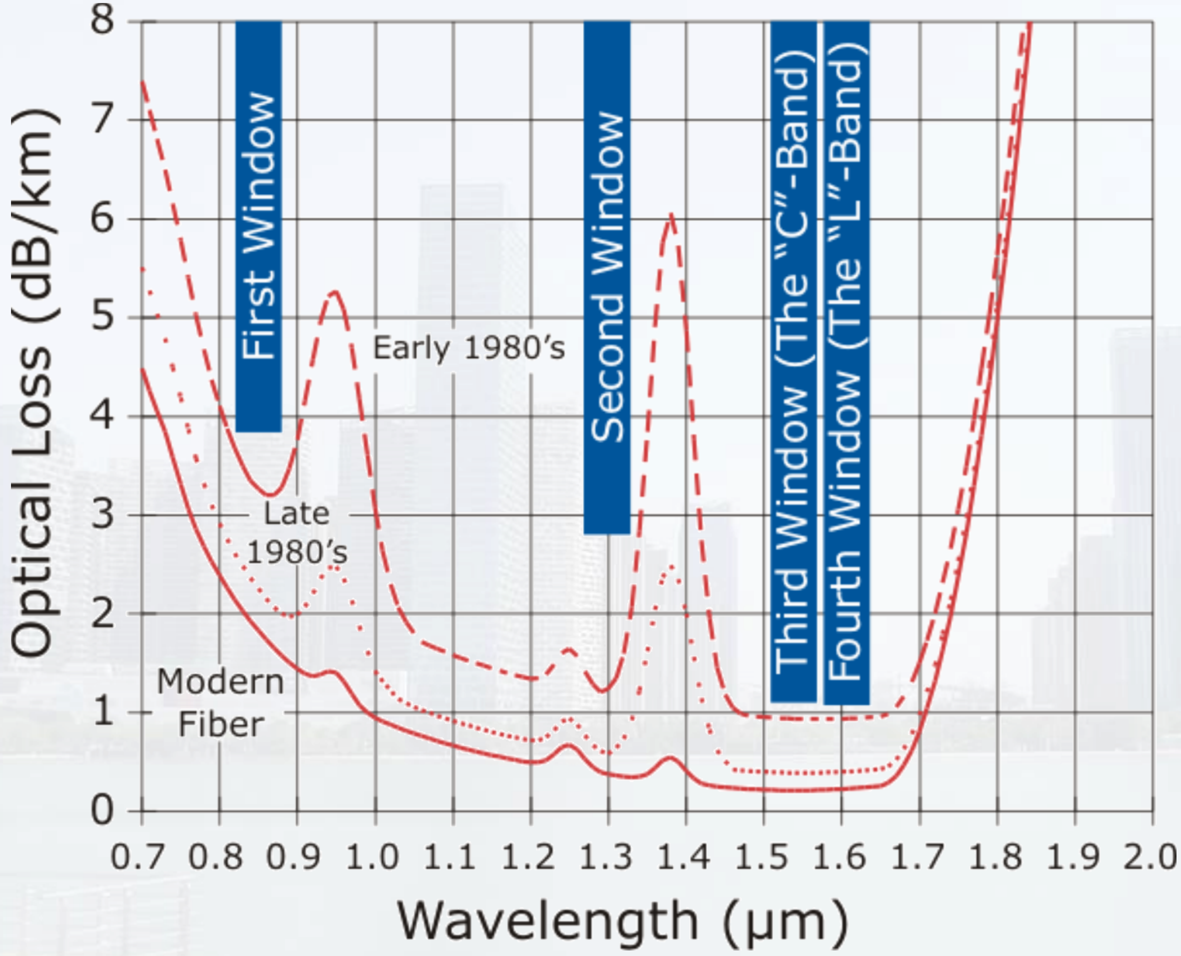
- Receiver tracks phase change in analog optical signal (“phase coherence”)
  - Dedicated oscillator on receive side of channel
  - Difference between reference and received signal to compute phase
  - Requires digital signal processing (“DSP”) on the receiver
    - Expensive, power-hungry, hard to integrate
    - Chromatic Dispersion Equalization to compensate for CMD
  - Also marketed as “coherent optics”

Quadrature Amplitude Modulation (QAM)

- Combine ASK and PSK
  - Multiple amplitude-modulated carriers
  - Carriers have different phase (e.g. sine and cosine)
- Heavy digital signal processing required



Optical fiber transmission bands

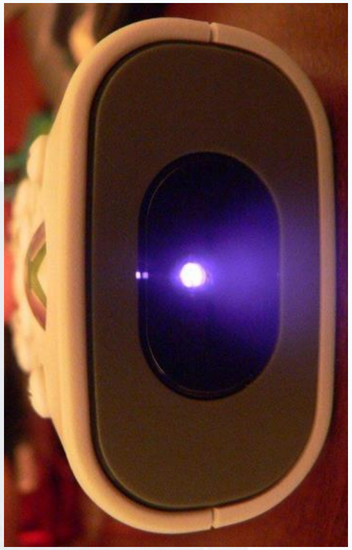


OM3, OM4, OM5 are usually specced for 850-953 nm.

O band ("second window") is mainly used for high-speed Ethernet transmission and Passive-Optical Network (PON) systems.

Optical fiber transmission bands

- "First Window" – 850 nm
  - Highest attenuation, "short range" O(100 m)
- "Second Window" – 1310 nm
  - "O-band", kilometer range
- "Third Window" – 1550 nm
  - "C-band", "Conventional band", long-reach DWDM
- "Fourth Window" – 1590 nm
  - "L-band", "Long band"
- The human eye can only see between 390 – 750nm
  - One trick to check for light in a fiber is to hold it up to your camera phone



Optical attenuation

- As light propagates in a fiber, its intensity decreases
  - Scattered by defects in the glass
  - Absorbed by impurities and converted to heat
- Attenuation is measured in “Decibels” (10 x log10(A/B))
  - 1/10 = 90% loss = -10 dB
  - 1/100 = 99% loss = -20 dB
- Optical power often expressed in “dBm”
  - 0.5 mW = -3 dBm
  - 1 mW = 0 dBm
  - 2 mW = 3 dBm
  - 100 mW = 20dBm
- Combining losses/gains expressed in dB = addition

Other causes for attenuation

- Insertion loss
- Return loss
- Bending loss
- Environmental contamination
- Radiation induced “darkening”

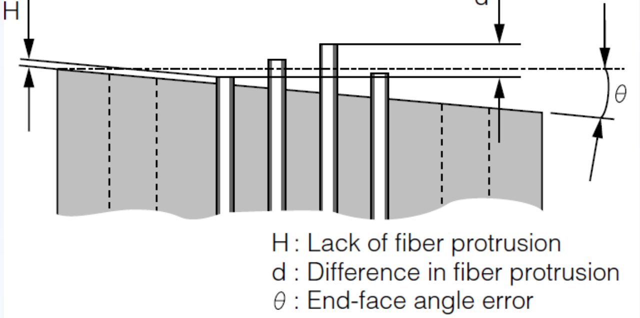
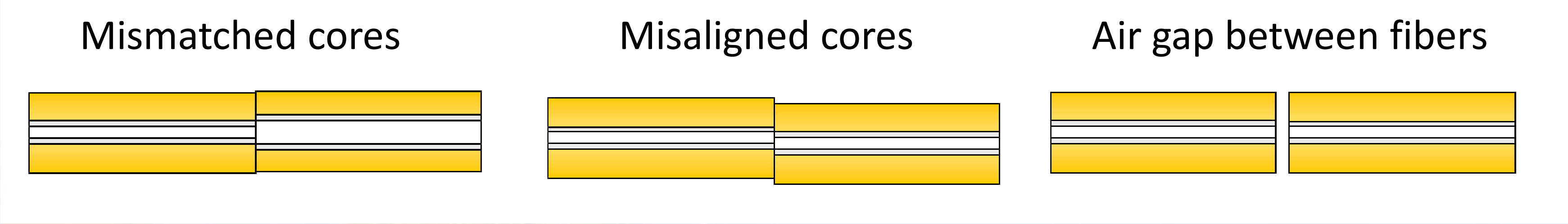


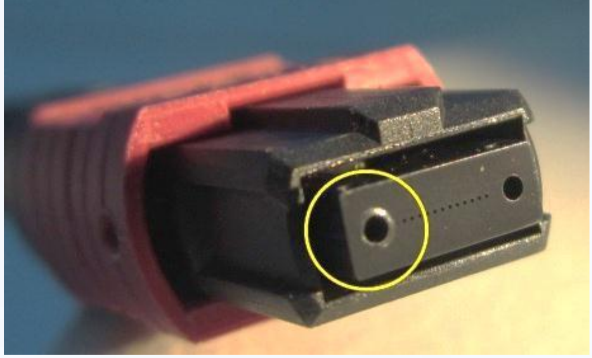
Fig. 9. Factors Affecting Physical Contact.

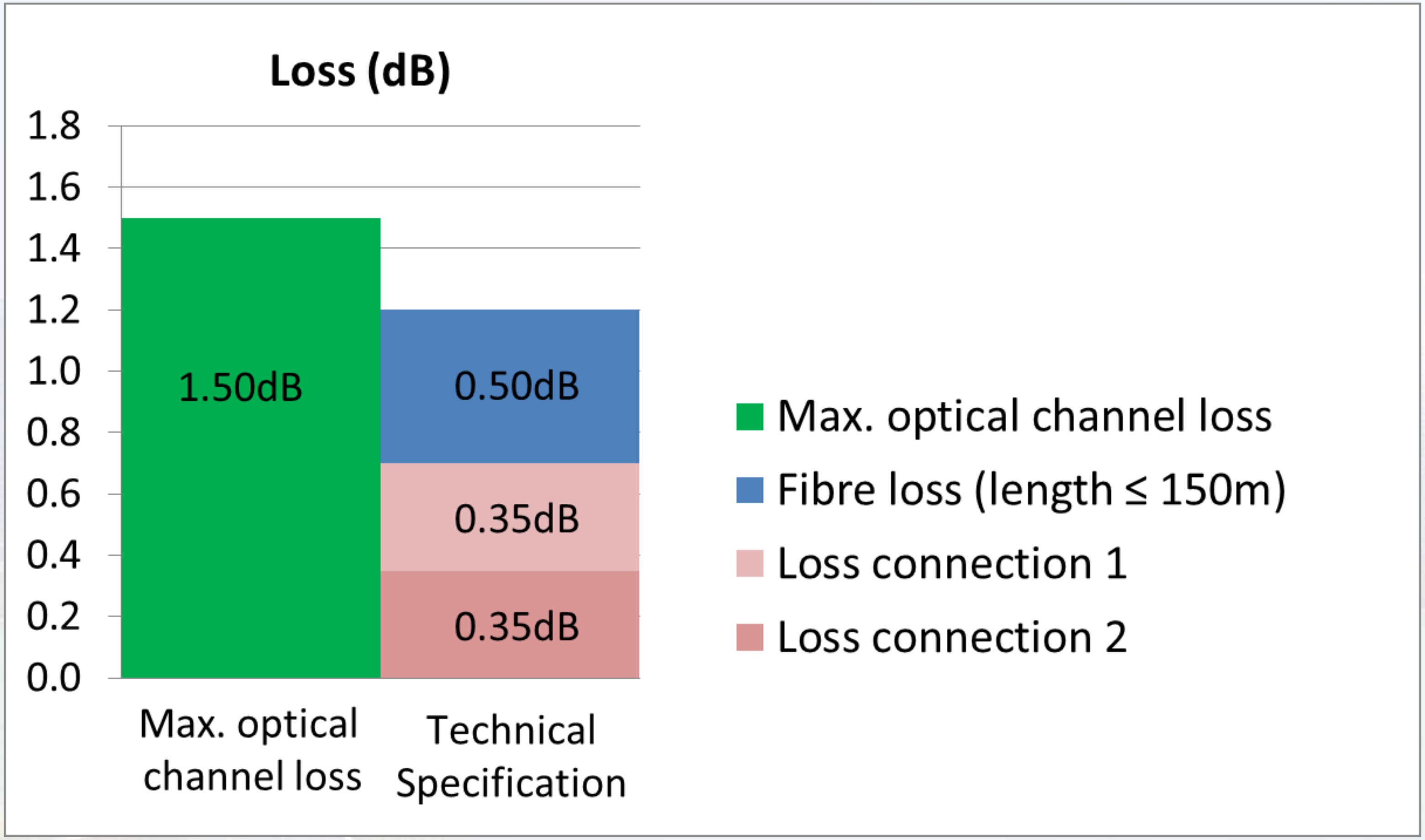
Insertion loss

- Number and quality of connectors and splices
- Standard < 0.60 dB (typ. 0.20 dB)
- «Elite» < 0.35 dB (typ. 0.10 dB)
- 40G OM3: 1.5dB total connector loss budget
- 40G OM4: 1.0dB total connector loss budget



Connector loss budget

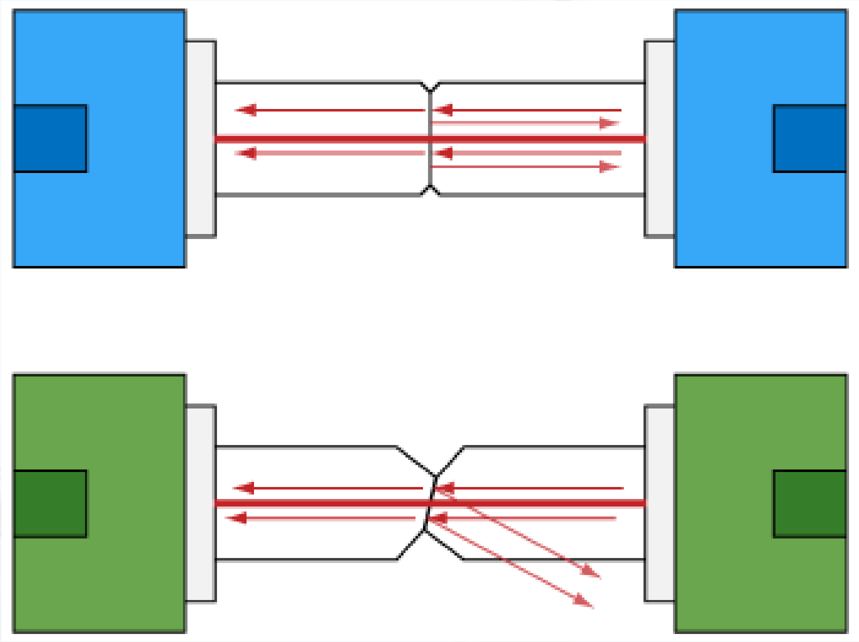




$$\mathrm{ORL(dB)} = 10\log_{10}\frac{P_\mathrm{i}}{P_\mathrm{r}}$$

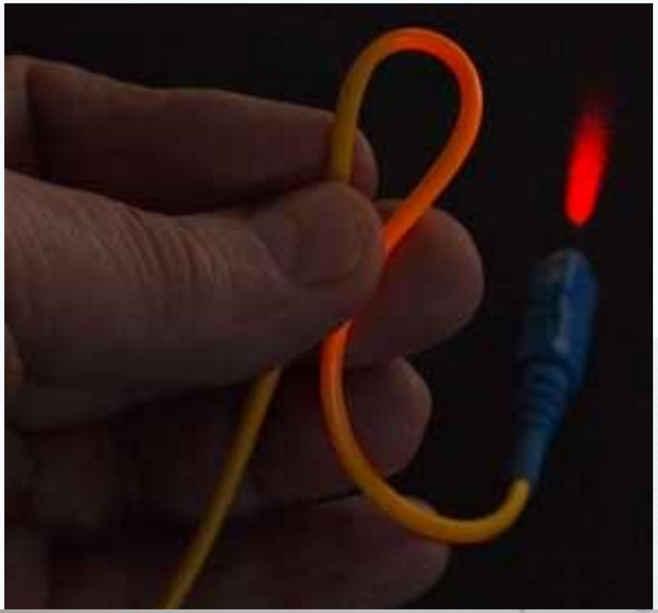
Return loss

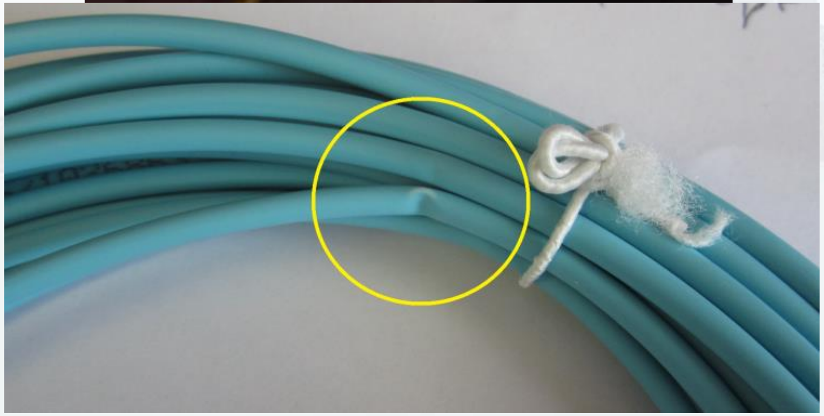
- Incident / reflected power [dB]
- (Ultra) Physical Contact (UPC) blue
  - (< -30 dB) < -55 dB ORL
- Angled Physical Contact (APC) green
  - 8° angle
  - < -65 dB ORL
  - For high-power emitters
- Are incompatible!



Bending loss

- Total internal reflection depends on critical angle
- Small bend radius causes light leaks
- Bend-insensitive fibers exist
  - Higher *n* cladding
  - Short range (patch panels)
- Exceeding the safe bend radius (kinking) <u>physically damages the fiber</u>

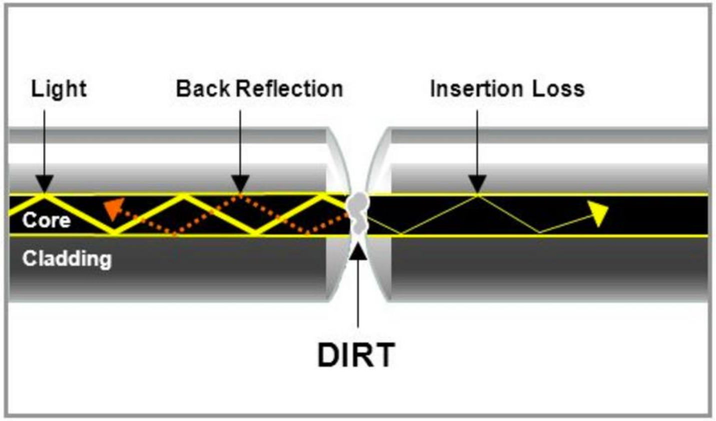


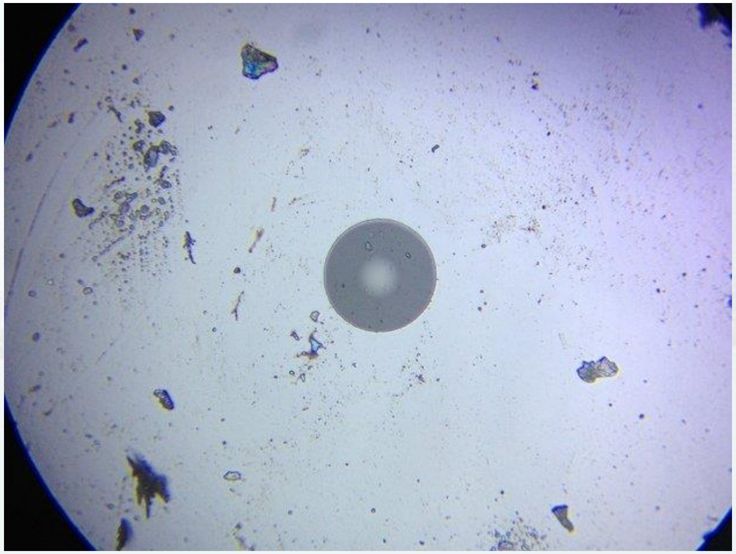


Dirt induced loss

- A µm speck of dust can completely (or partially) cover the core
- Insertion force can dent the optical interface
- Don't leave fibers exposed
- Clean **both sides** before mating
  - Cleaning cassettes
  - Insertable tape rolls

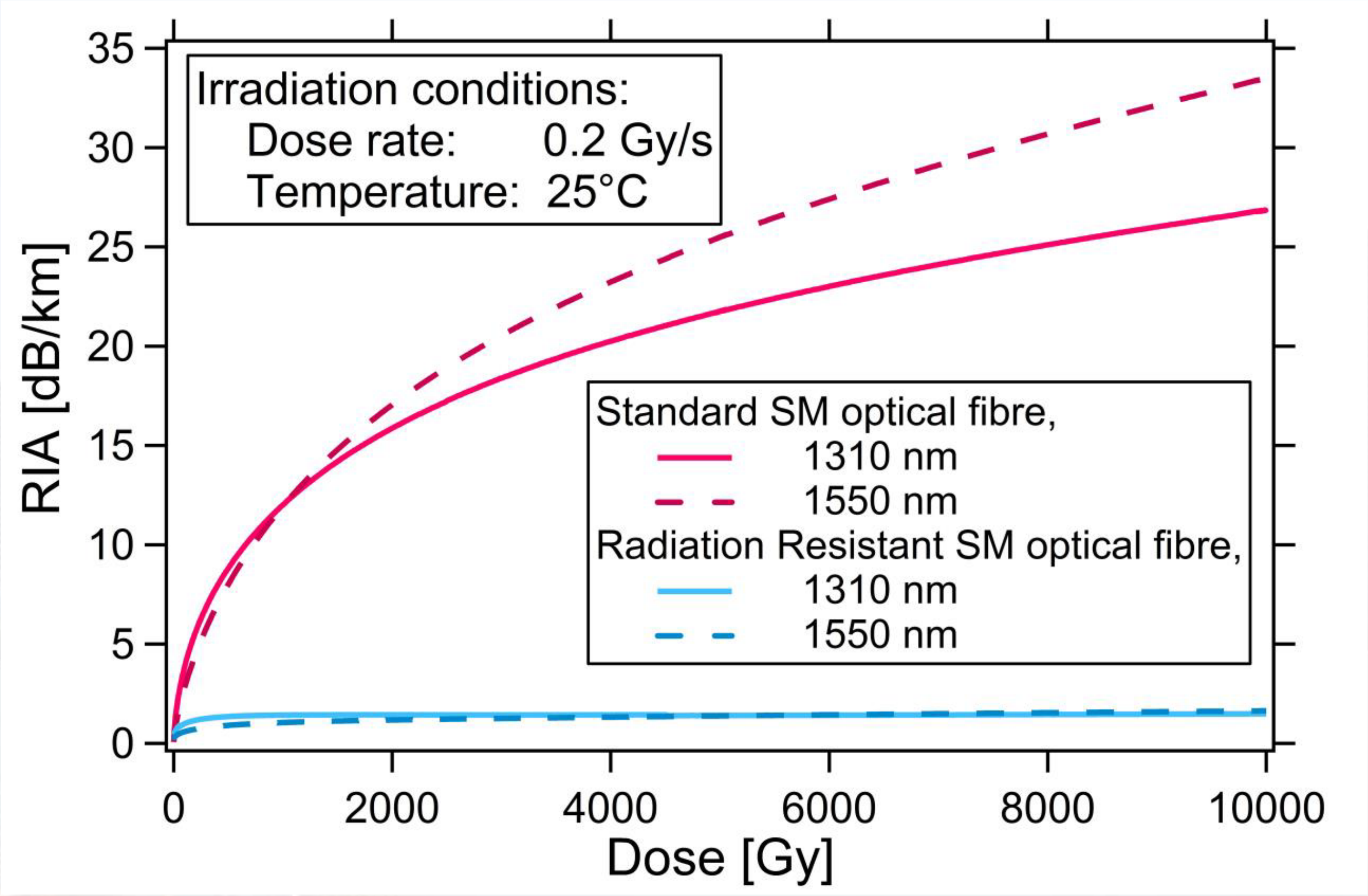
Optical fiber visual inspection





Radiation induced loss

Use rad-hard fibers with special doping and manufacturing processes

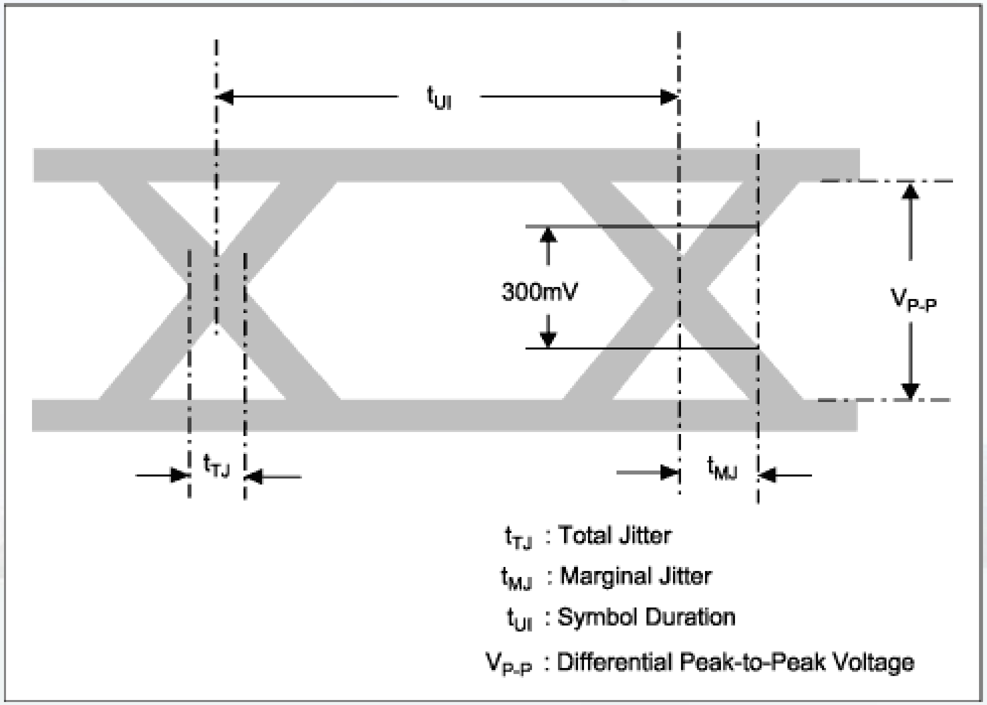


How to compensate for attenuation/distortion

- On the transmitter
- On the receiver
- Forward error correction

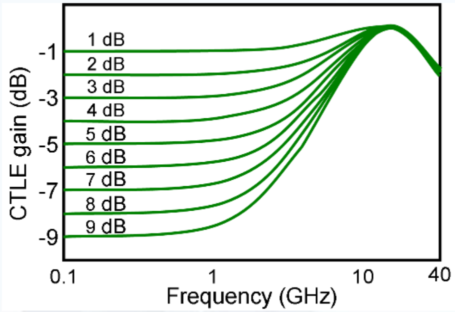
Eye diagram

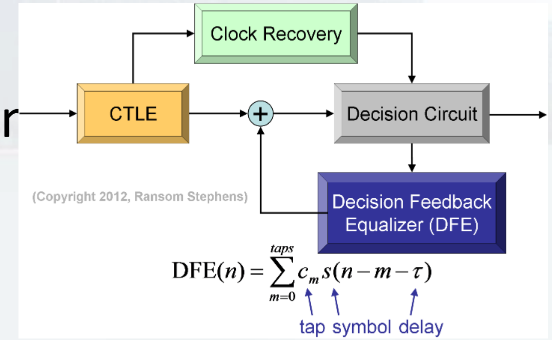
- I→V (trans-impedance amplifier)
- Superimpose all possible binary transitions
- Quickly quantify noise and inter-symbol interference (ISI) effects
- Many standards mandate minimal eye parameters



Signal optimization techniques

- Pre-emphasis (already used at 10G)
- Tx FFE (Feed Forward Equalization)
  - Emphasis of high-speed content to combat frequency dependent attenuation in channel
- Rx CTLE (Continuous Time Linear Equalization)
  - De-emphasizes low frequency, peak at Nyquist, filters off high frequency
- Rx CDR (Clock and Data recovery)
  - Used together with CTLE or other equalization
- Rx DFE (Decision Feedback Equalization)
  - Non-linear, feedback from decision circuit
  - Power hungry…

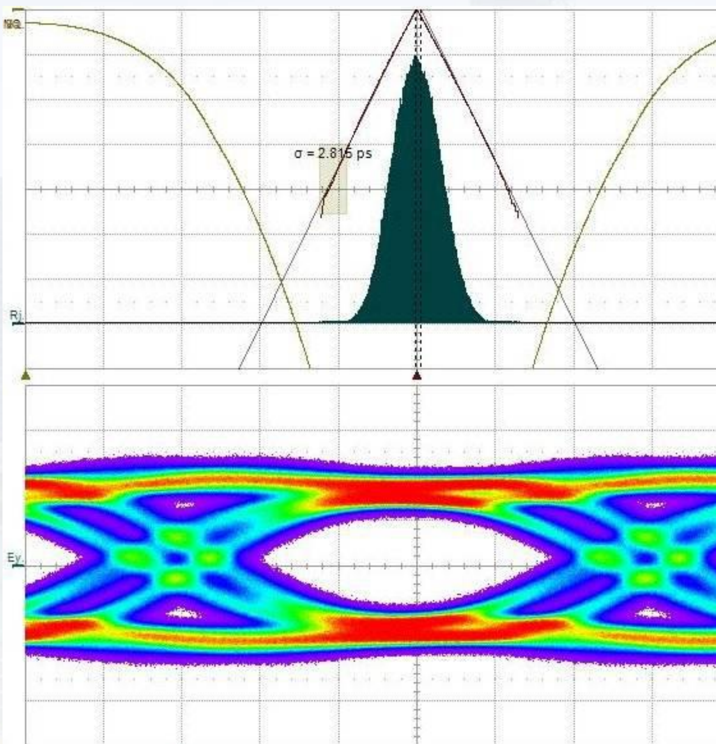




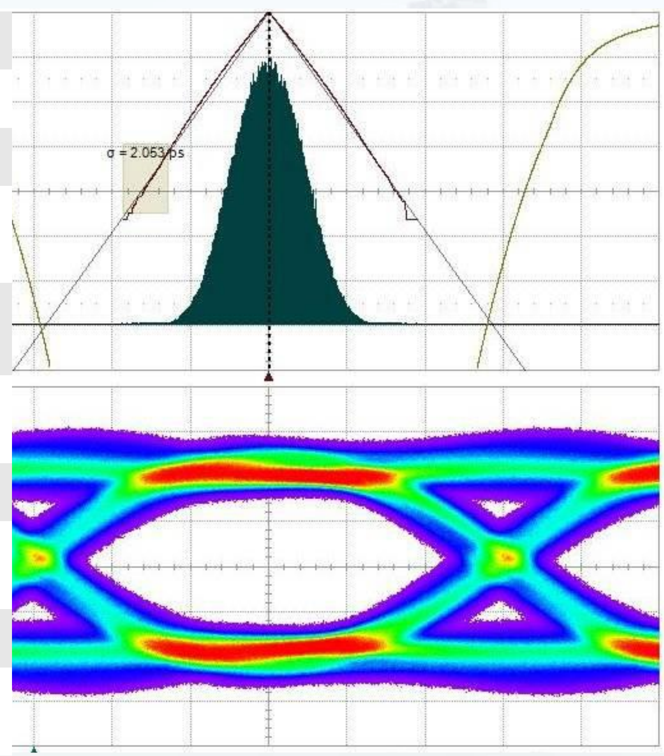
Example: 10GBASE-R

No de-emphasis nor equalization on optical drivers

With de-emphasis and equalization



	No de-emphasis nor equalization	With de-emphasis and equalization
Eye aperture	0.29	0.54
BER	$<10^{-15}$	$\ll 10^{-16}$
Tot. jitter @ 10^{-12}	62.7ps	39.3ps
Random jitter	3.13ps	2.25ps
ISI	25.5ps	13.6ps



https://indico.cern.ch/event/297003/

Forward Error Correction (FEC)

- Insert “redundant” information before transmission
- Receiver can detect if errors occurred, and where
- Receiver can computationally fix errors, up to a design limit dependent on the algorithm
- 1st generation FEC (Reed-Solomon)
  - 6% overhead for ~6 dB BER gain
- 2nd generation FEC (EFEC)
  - 7% overhead for 8~9 dB BER gain
- 3rd generation FEC (SD-FEC)
  - 20-25% overhead for 10~11 dB BER gain
- FEC is now **mandatory** in most high-speed standards, like 100GbE and beyond

Other definitions

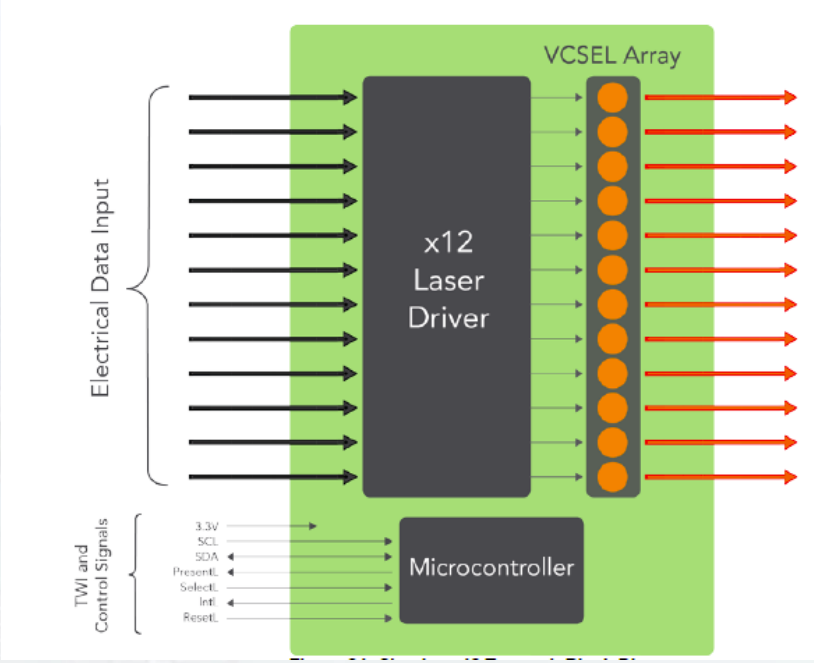
- Transmitter launch power
  - dBm power @ λ
- Optical Modulation Amplitude (OMA)
  - $\mathrm{OMA} = P_1 - P_0$
- Extinction ratio
  - $r_e = \frac{P_1}{P_0}$
- Receiver sensitivity
  - Min. received power @ BER target
- Dynamic range
  - (Worst-case max input power - Sensitivity)

Outline

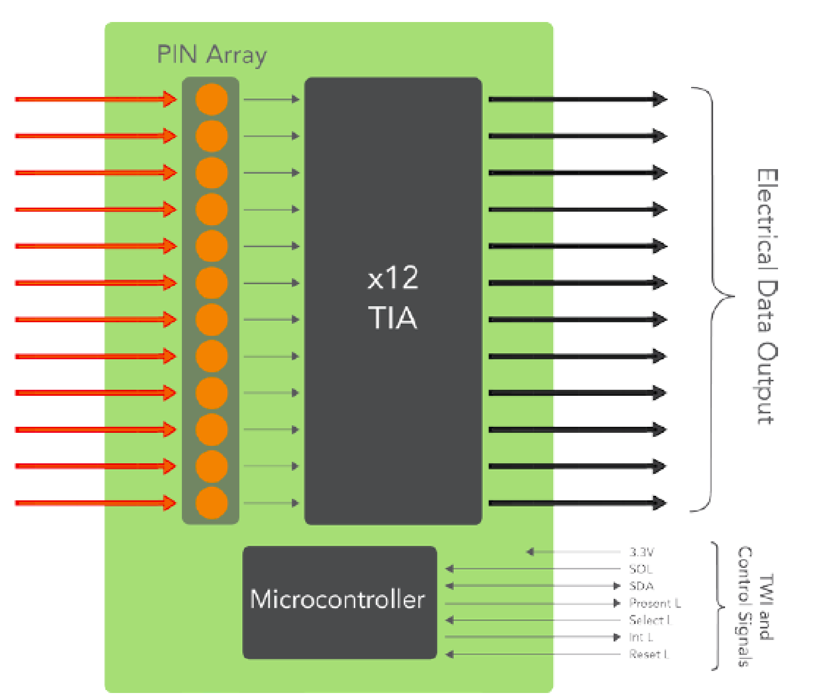
- Fundamental concepts
- <u>**Physical devices**</u>
- Applications
  - Frontend optical links
  - Backend optical links
- Installation / Commissioning

Vertical-Cavity Surface-Emitting Laser (VCSEL)

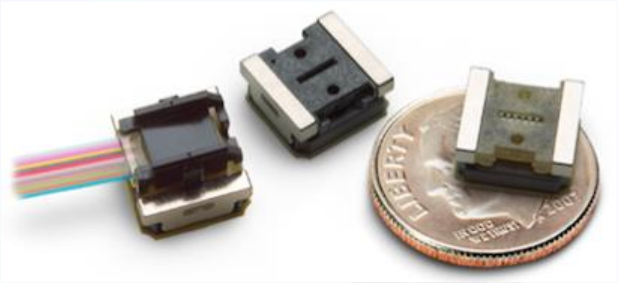
Tx



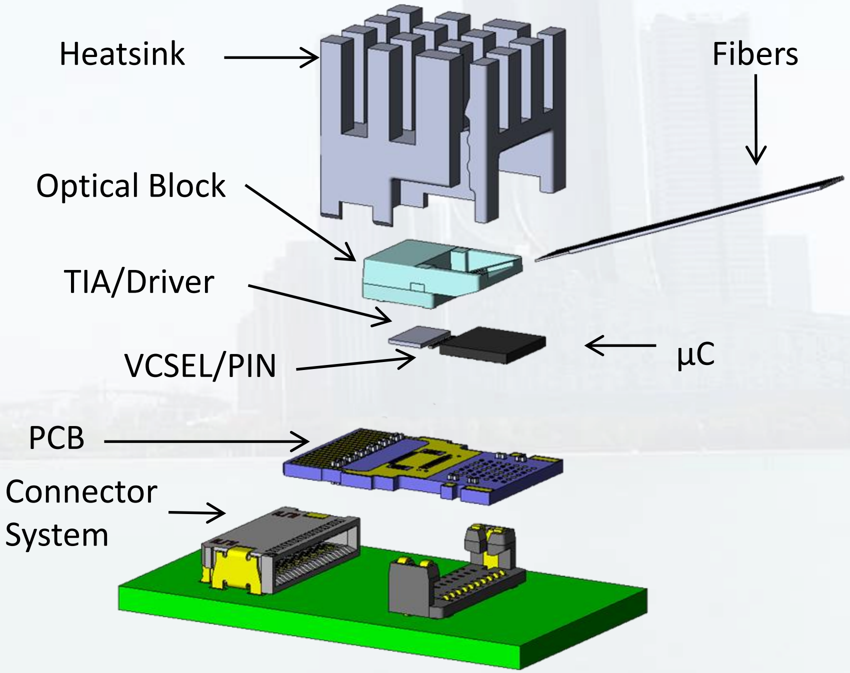
Rx



VCSEL optical engine assembly

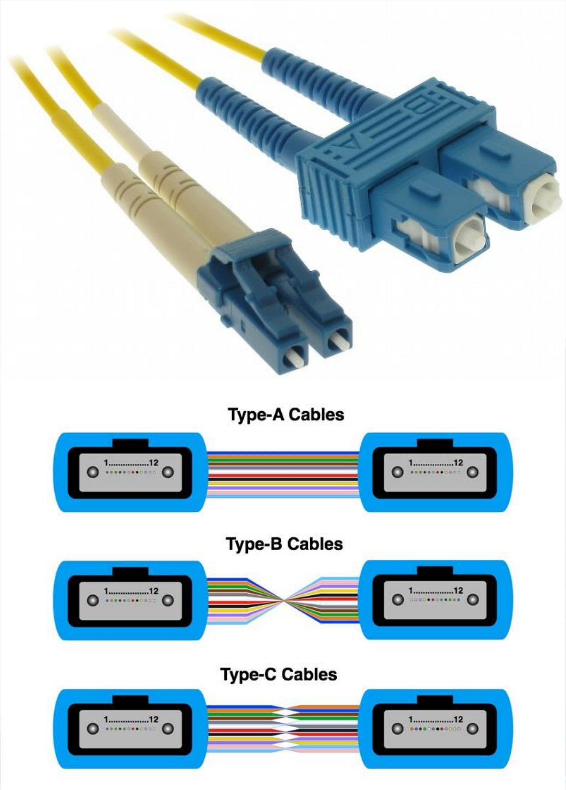


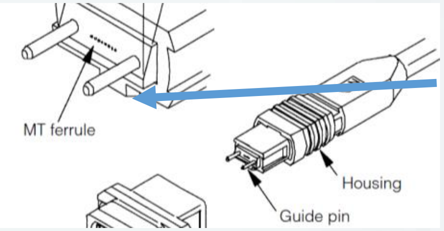
- Can be integrated with other circuitry
- Can be packed in dense 1D and 2D arrays
- Can be tested pre-dicing
- High quantum efficiency -> low power
- Good optical coupling via MO interfaces (e.g. PRIZM® LightTurn®)



Optical connector assemblies

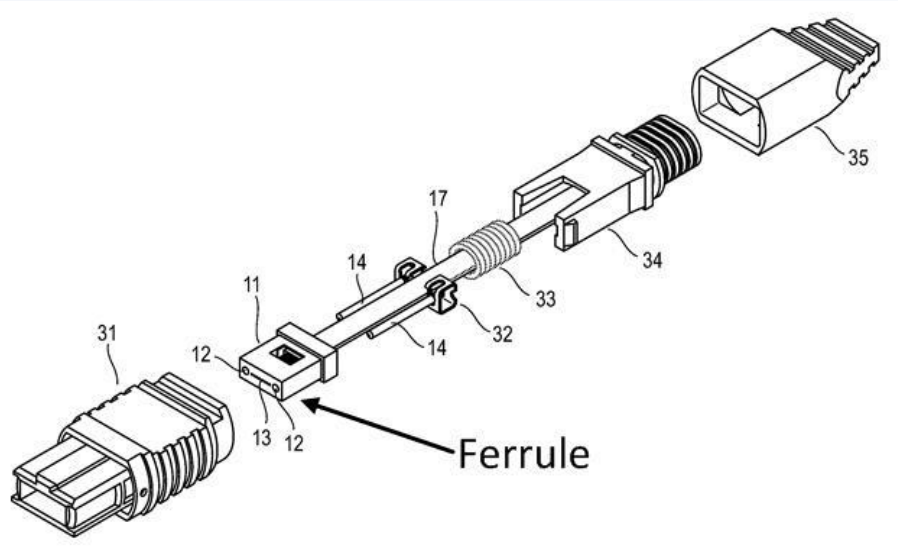
- Form factors
  - SC (x1, x2, obsolete)
  - LC (x1, x2)
  - MPO (12, 16, 24, 32, 48)
  - MTP (high-quality MPO)
- Polarity
  - Straight (A)
  - Crossover (B)
  - Flipped-pairs (C)
- Gender
  - Male
  - Female

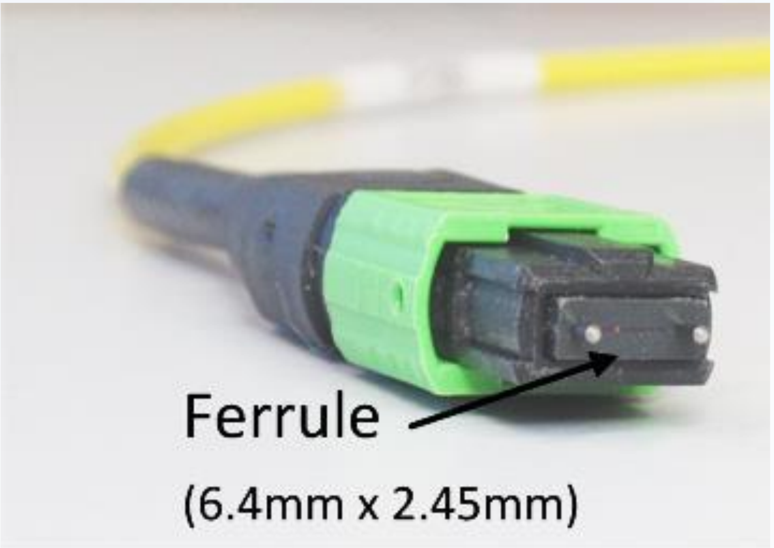


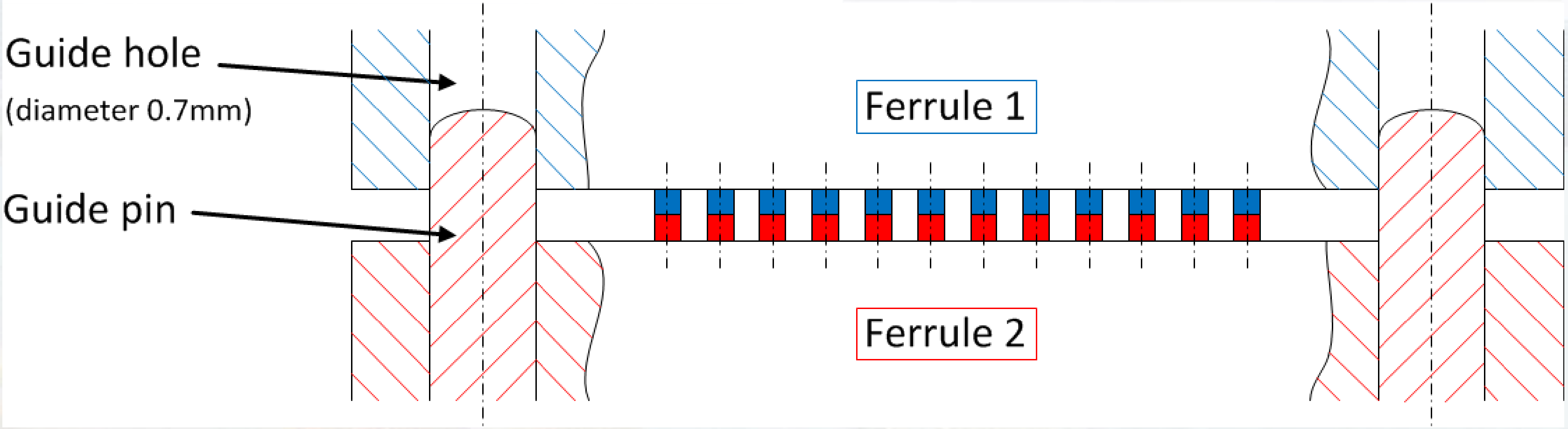


key

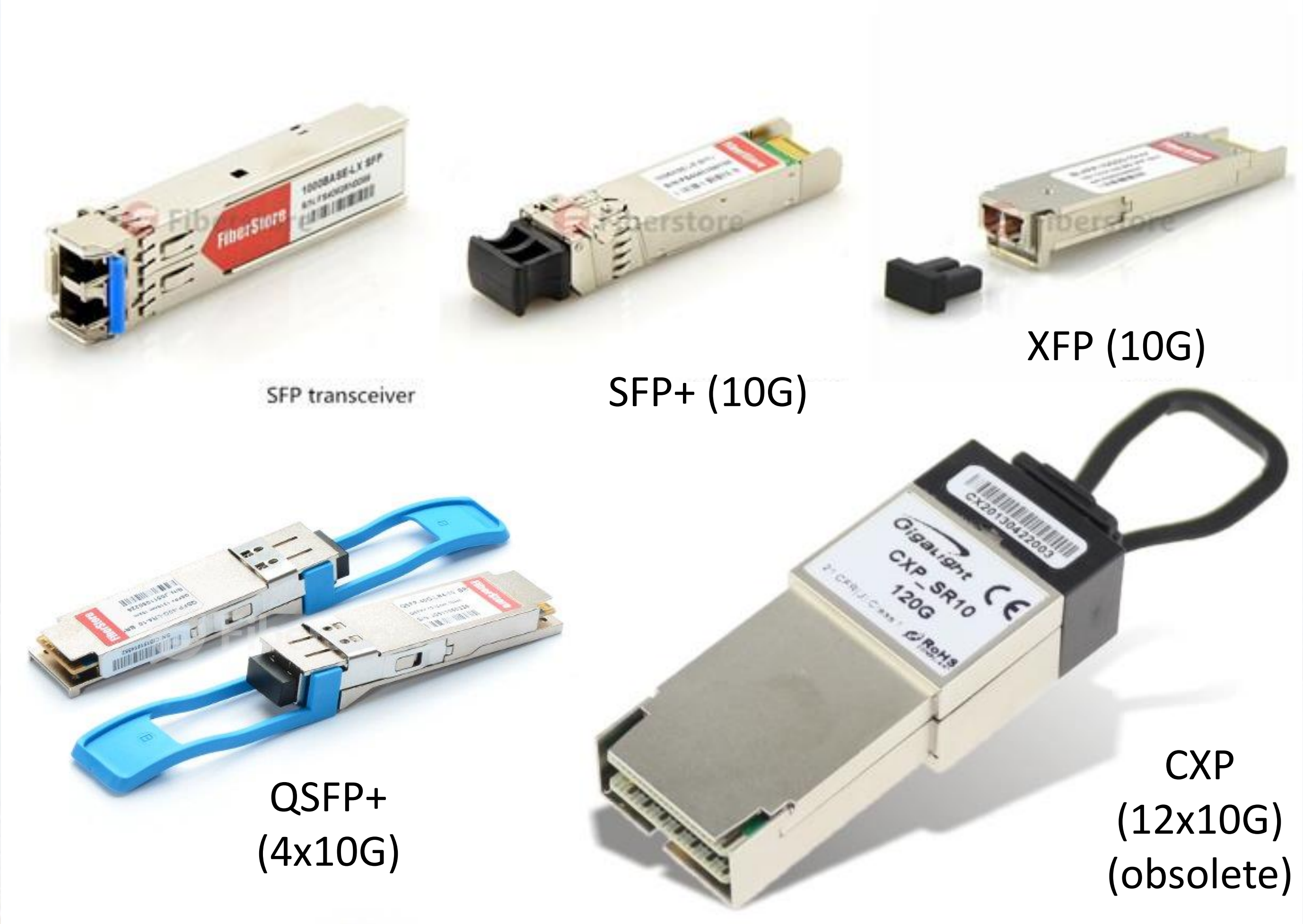
MTP/MPO optical assembly



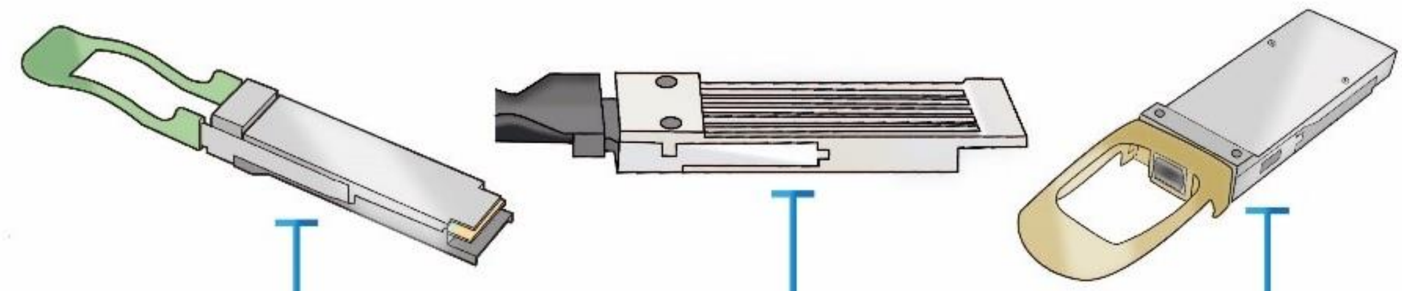




Pluggable form factors ("slow")



Pluggable form factors ("fast")

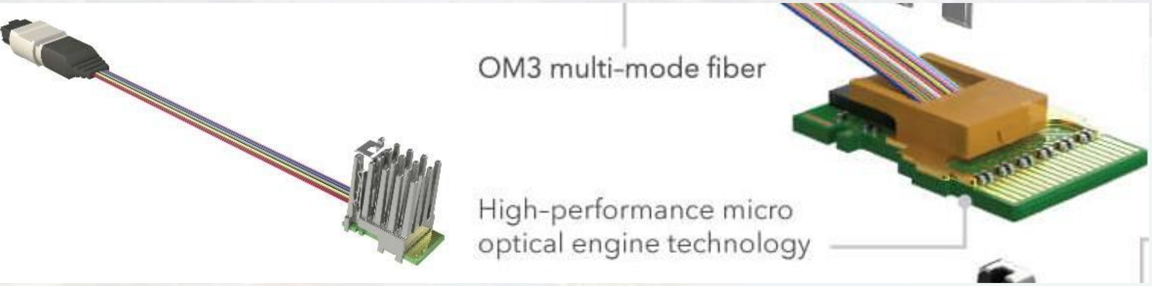


	QSFP56-DD	OSFP	CFP8
Power consumption:	max 12W	max 15W	max 24W
Backward compatibility with QSFP+, QSFP28 and QSFP56:	yes	Through adapter	No
Size:	Similar to QSFP28	Slightly bigger than a QSFP28	Slightly smaller than CFP2
Thermal management:	Indirect	Direct	Indirect

Proprietary packages (high density)

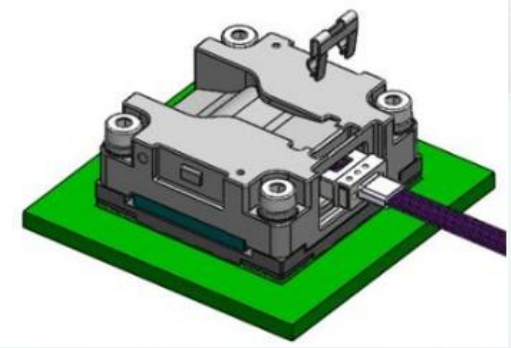
Samtec FireFly ECUO

- 70m OM3 850nm
- 12 simplex or 4 duplex
- Ribbon MPO-12 or MPO-24
- 3.3V (+1.8V si duplex 28G)
- 14 / 25 / 28 Gb/s
- 240mm2 footprint
- T12 | R12 | B04 variants (14G)
- B04 variant (28G)



III-VI Finisar BOA

- 70m OM4 850nm (possible 70m OM3)
- 12 duplex
- MT-24 connector (no ribbon)
- 6W (2.5V + 3.3V rails)
- 25Gb/s
- 625mm2 footprint
- CFP MSA slow control interface



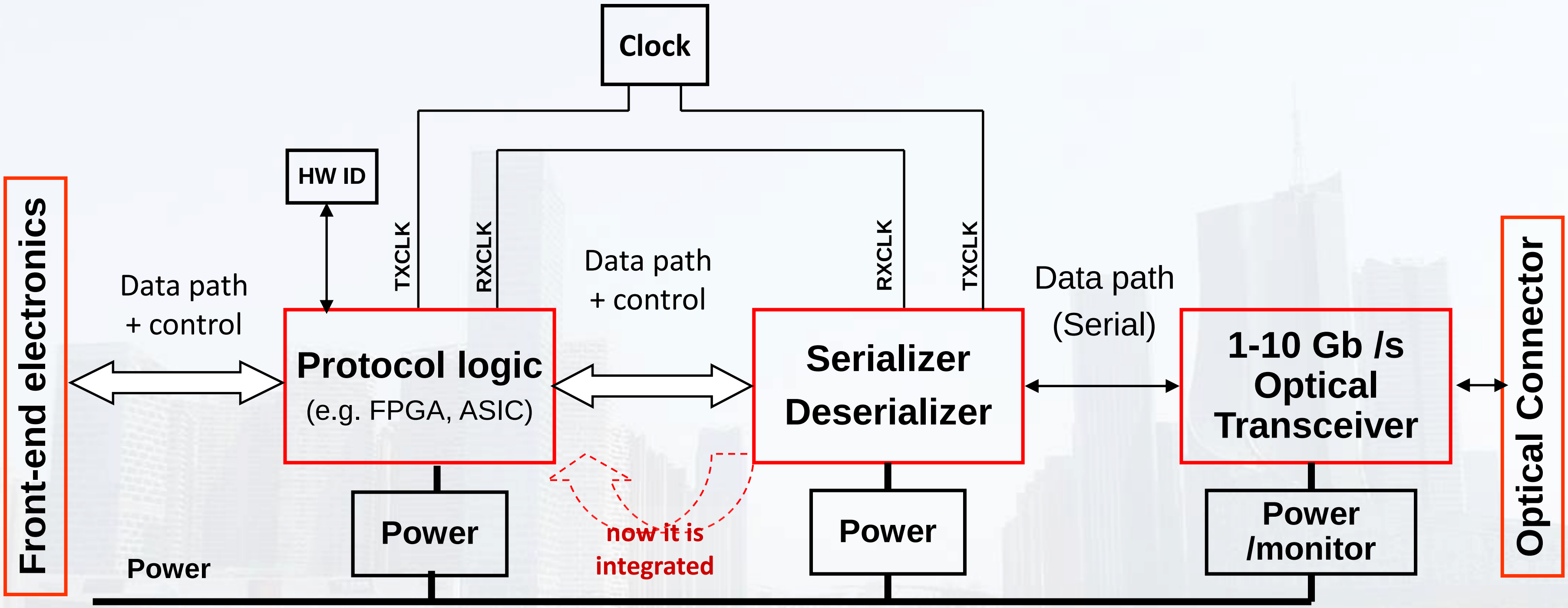
Active Optical Cables (AOC)

- Great for point to point connections, when length is known in advance
- Handled (almost) like a copper cable, no optical connector or cleanliness related problems
- Cheaper than discrete optics
- Available in “Octopus” breakouts

Outline

- Fundamental concepts
- Physical devices
- **<u>Applications</u>**
  - Frontend optical links
  - Backend optical links
- Installation / Commissioning

System integration

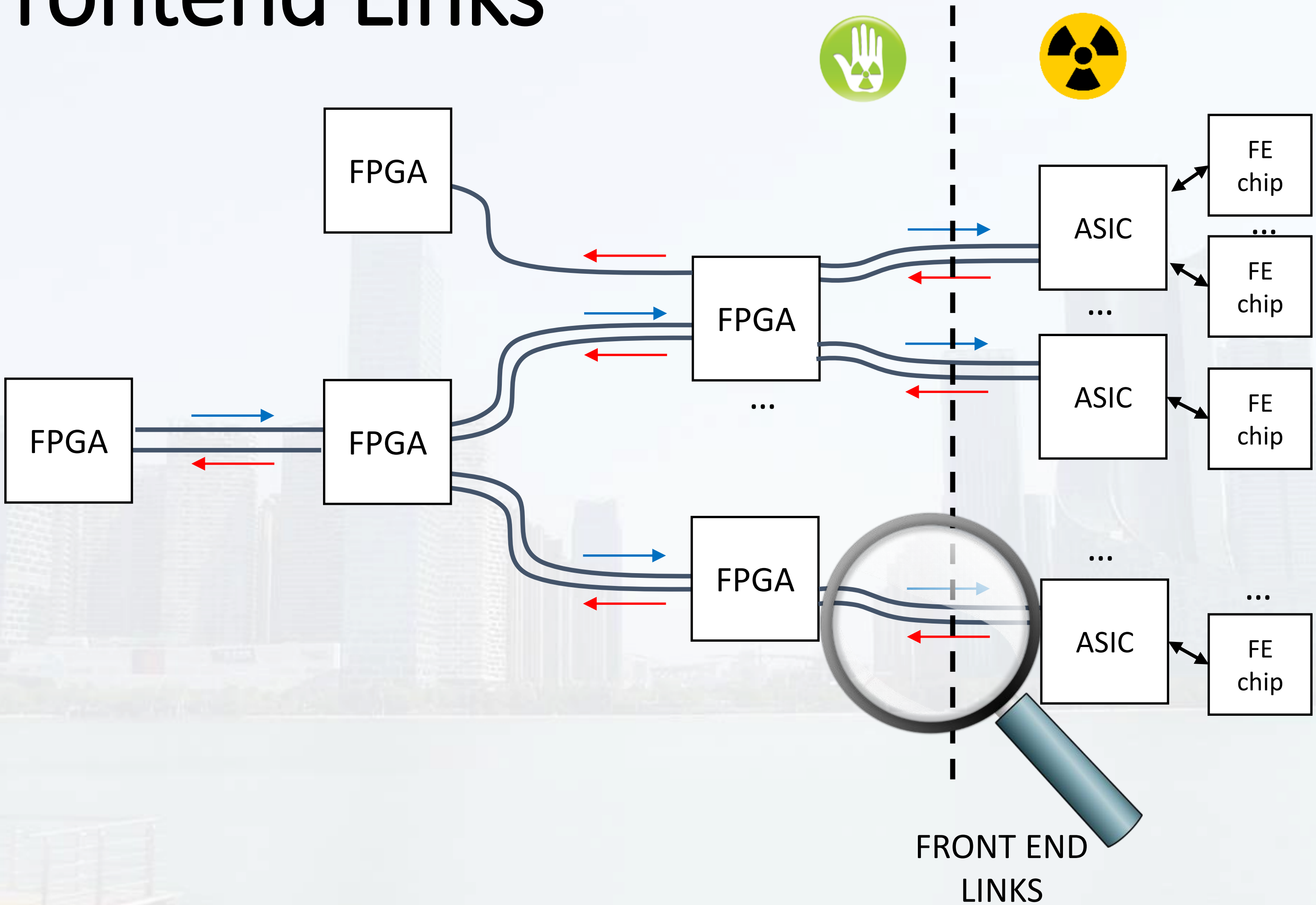


- **FIFO like parallel interface**
- **Simplex, half-duplex, or duplex**

- **Physical Coding (e.g. 8B/10B)**
- **Link Control, Data Framing,**
- **Flow Control**
- **Error Detection**

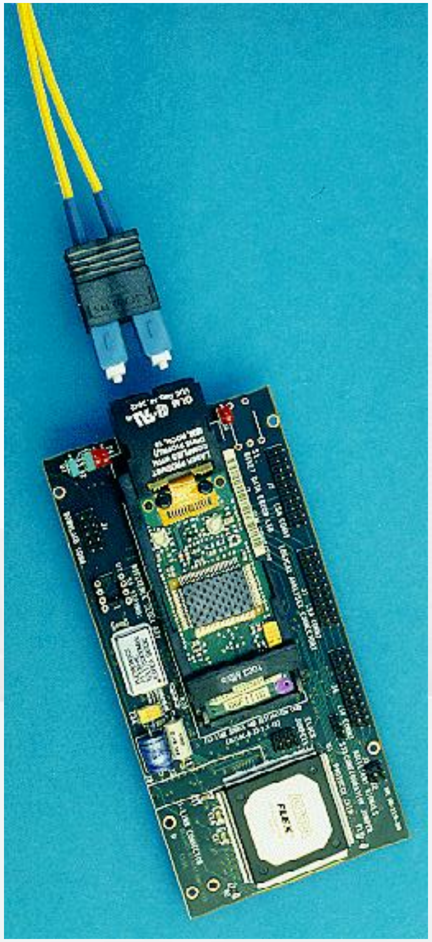
- **E/O conversion**

Frontend Links



Frontend Links: S-LINK

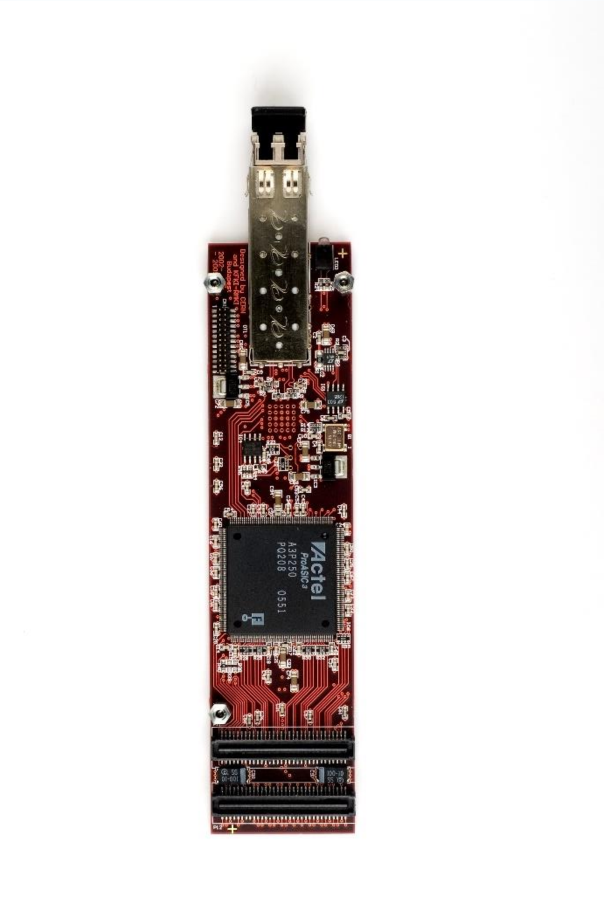
- Defined in 1995
- Duplex
- PCI interface
- 32-bit: ≤ 132 MB/s
- 8-bit: ≤ 64 MB/s
- Dual SC
- MMF
  - 50/125: 2 km
  - 62.5/125: 0.7 km
- OPB: 0.25 dB



http://cern.ch/HSI/**s-link**

Frontend Links: DDL

- Defined in 1997
- Duplex
- PCI Interface (RORC)
- 2.125 Gb/s (max. 2.5 Gb/s)
- 850 nm
- MMF: 50µm or 65µm
- Up to 500m



http://cern.ch/ddl

Frontend Links: Versatile Link

https://espace.cern.ch/project-versatile-link/public/default.aspx

- **3.2 Gbit/s** or **4.48 Gbits/s** user bandwidth
- Optional FEC
- SFP-like form factor
- Deterministic latency in both directions
- Radiation hard – qualified for:
  - 1 MRad total dose
  - 5e14 neq/cm²
- FE interface: 10 to 40 **E-links:** SLVS based with 320, 160 or 80 Mbit/s
- “Low”-power
  - <1.5W, 2.2W Worst-case

Frontend Links for LHC Phase-2 upgrades

Two partner projects

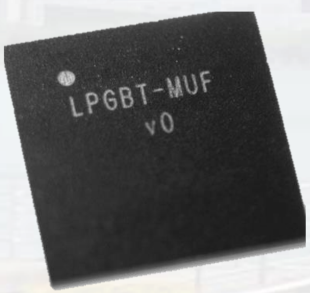
- **Versatile Link+** Optical Link: Custom module, fibre plant, commercial module
- **lpGBT** Chipset & protocol development: SerDes, LDD, TIA, FPGA core

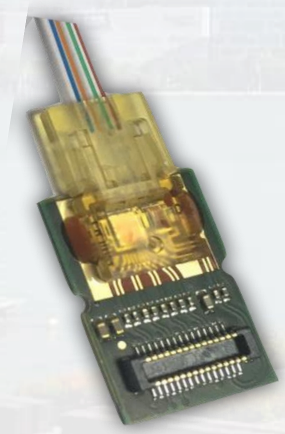
Asymmetric Data-rates

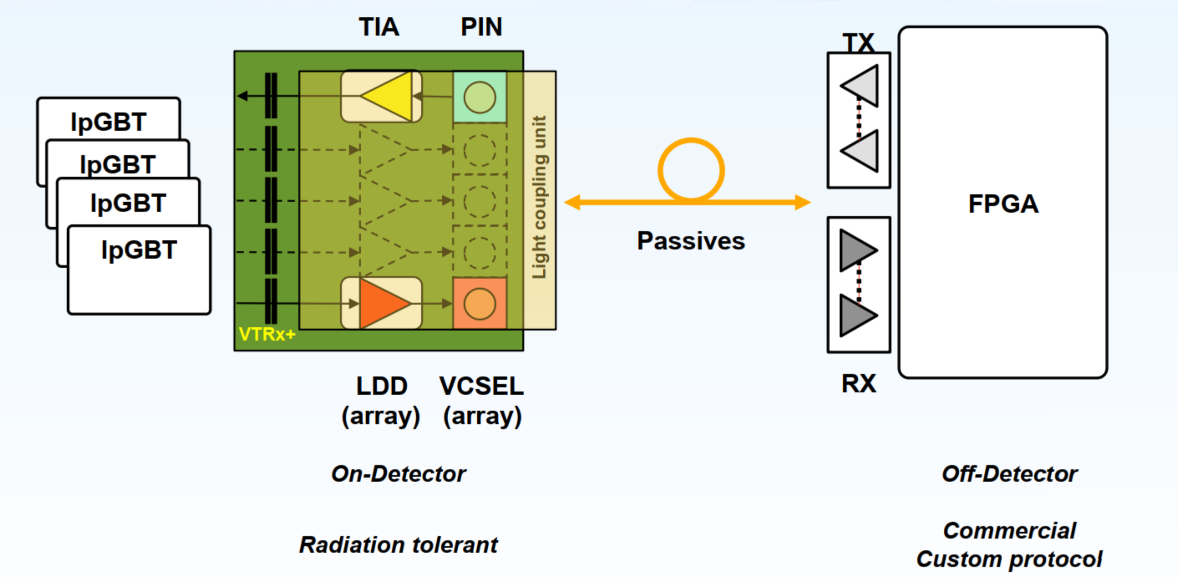
- 5 or 10 Gb/s upstream (out of detector)
- 2.5 Gb/s downstream

Harsh environment

- 1 MGy, 3x10^15 neq/cm2
- -35 to +60 °C







Versatile Link Plus – VTRX+

Miniaturized and pluggable
- Electrical connector
- Optical pigtail

Up to 4Tx+1Rx configuration
- configurable by channel masking

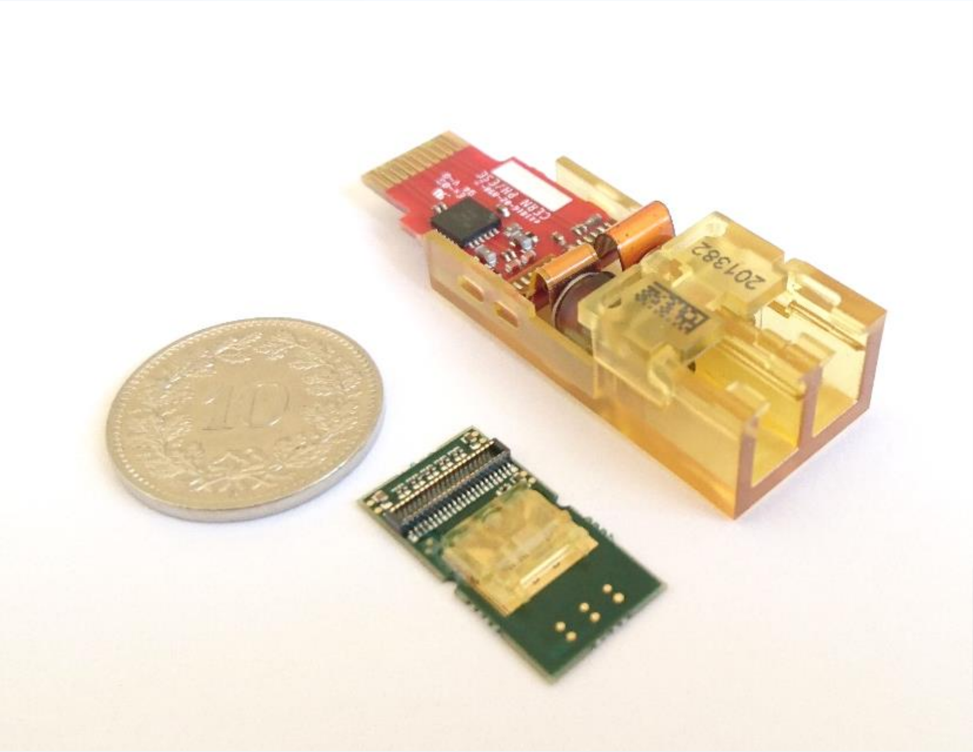
Transmitters
- 1×4 850 nm VCSEL array
- 5 and 10 Gb/s

Receiver
- InGaAs photodiode
- 2.5 Gb/s

For harsh environments
- Temperature: -35°C to +60°C
- Total dose: 1 MGy
- Total fluence up to 1×10^15 n/cm2 and 1×10^15 hadrons/cm2

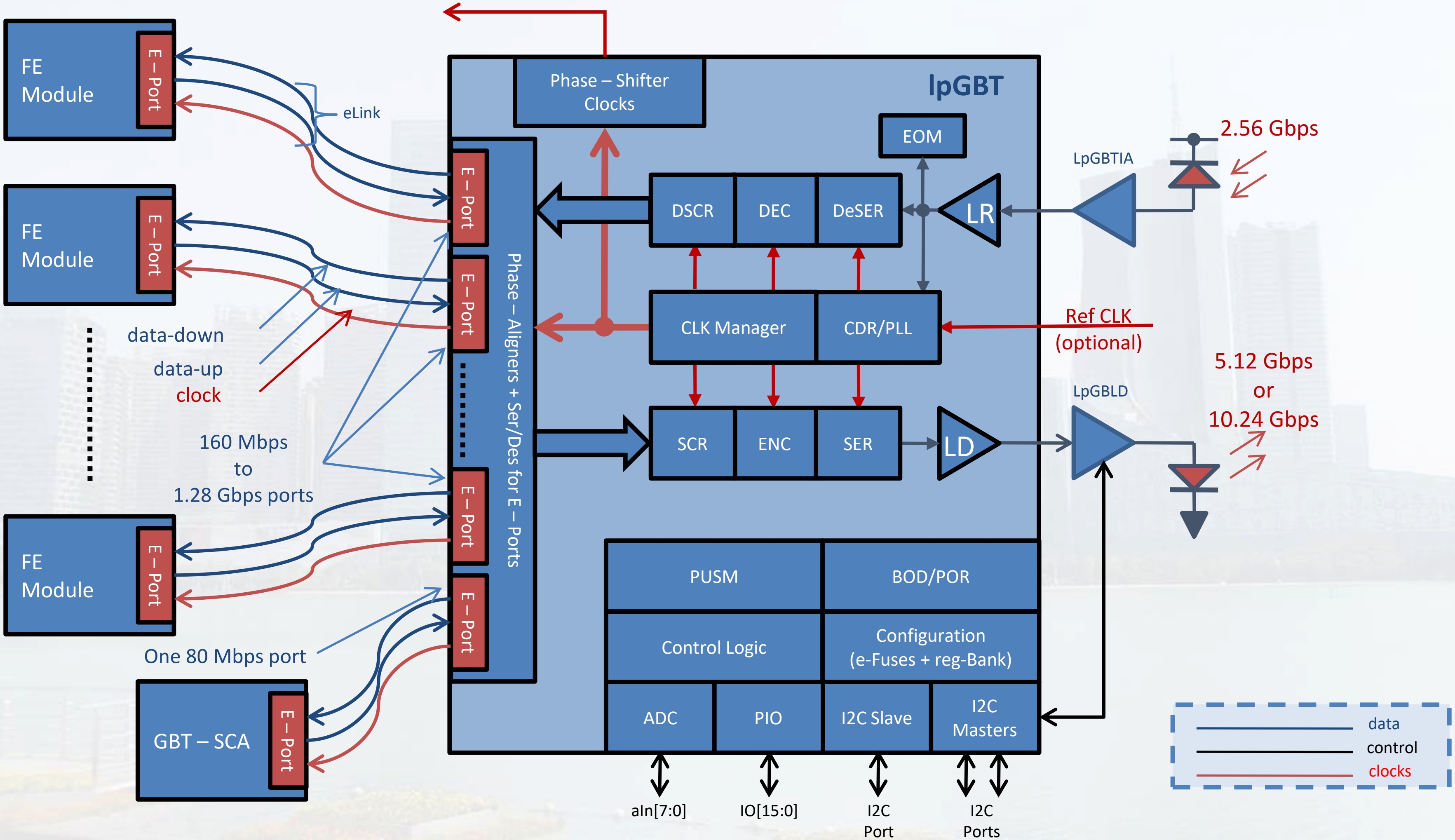
Recovered clock jitter
- ~ 5ps rms

Fixed and Deterministic Latency
- 1ps precision when Tclink is used

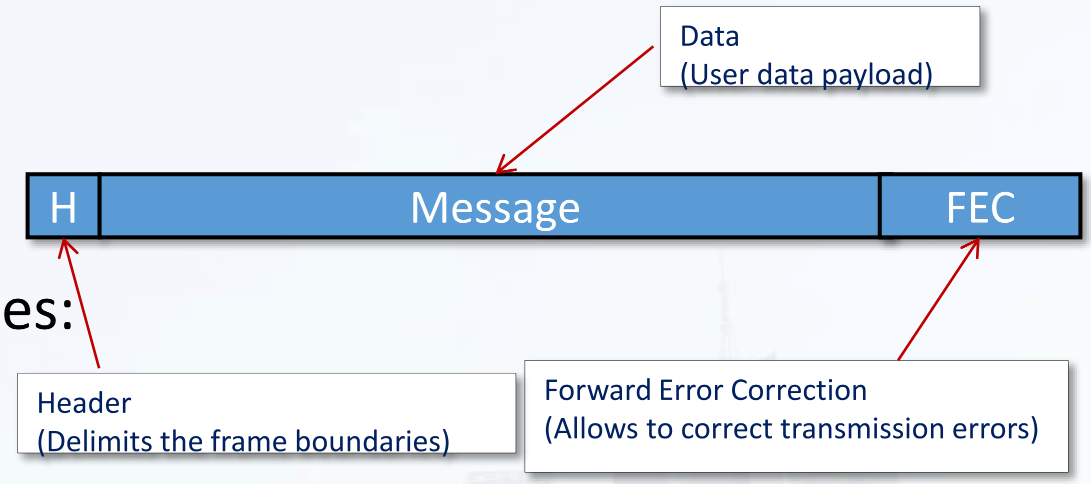


lpGBT system block diagram

https://espace.cern.ch/GBT-Project/LpGBT



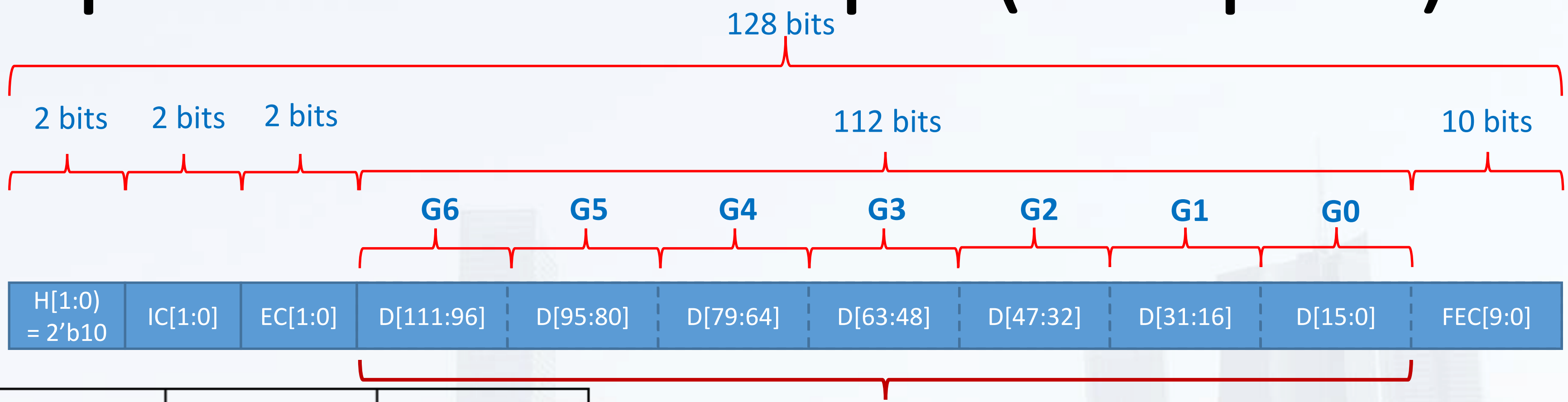
lpGBT uplink



- The LpGBT supports the following uplink data rates:
  - 5.12 / 10.24 Gbps
- Data is transmitted as a frame composed of:
  - Header
  - The data field
  - A forward error correction field:
    FEC5 / FEC12
- The data field is scrambled to allow for CDR operation at no [additional] bandwidth penalty

	uplink			
	5.12 Gbps		**10.24 Gbps**	
	FEC5	**FEC12**	**FEC5**	**FEC12**
Frame [bits]	128		256	
Header [bits]	2		2	
Data [bits]	116	102	232	204
FEC [bits]	10	24	20	48
Correction [bits]	5	12	10	24
Efficiency	91%	80%	91%	80%

lpGBT e-links example (5G uplink)



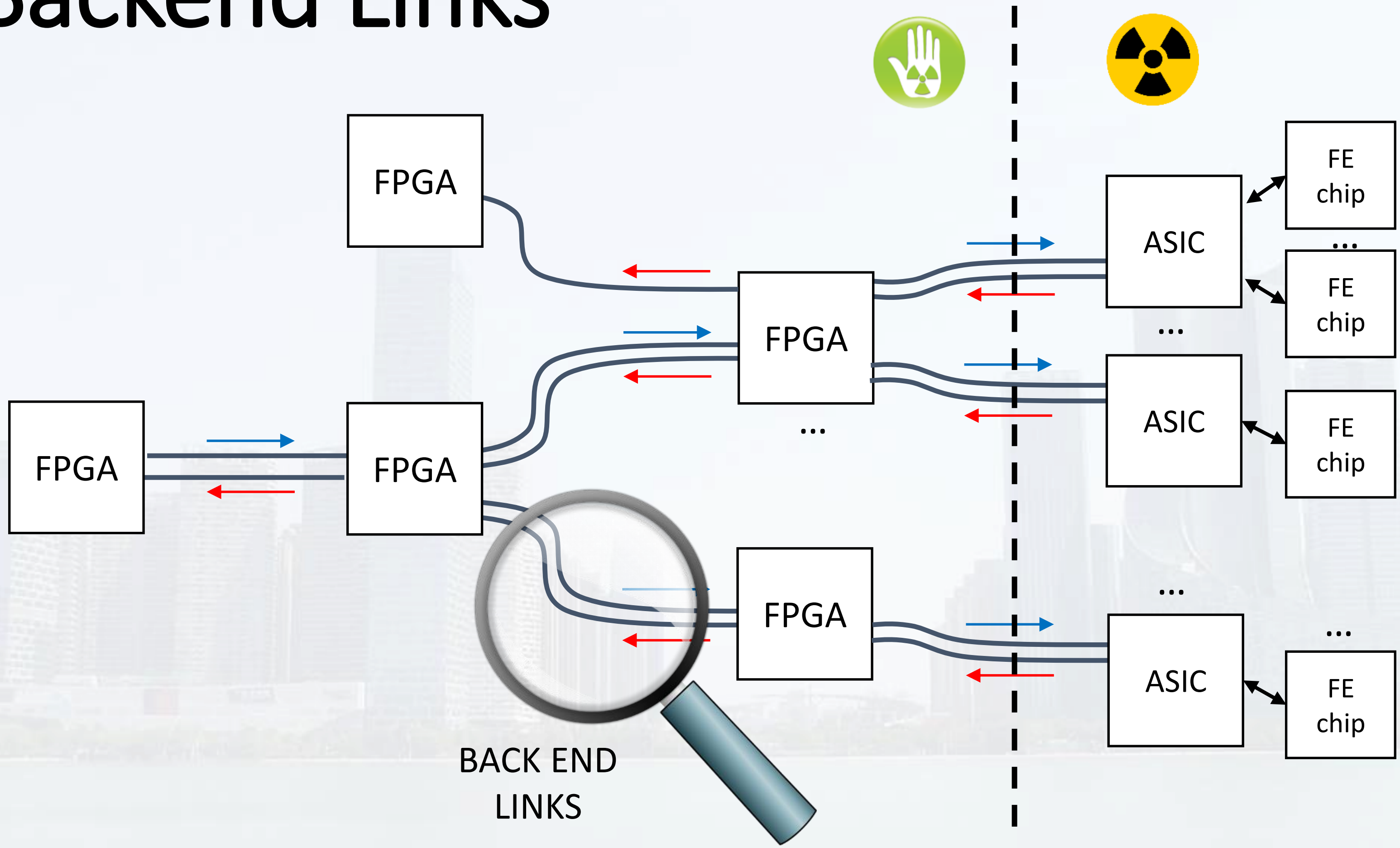
Frame	Function	I/O Group
FRMUP[9:0]	FEC[9:0]	
FRMUP[25:10]	Data[15:0]	0
FRMUP[41:26]	Data[31:16]	1
FRMUP[57:42]	Data[47:32]	2
FRMUP[73:58]	Data[63:48]	3
FRMUP[89:74]	Data[79:64]	4
FRMUP[105:90]	Data[95:80]	5
FRMUP[121:106]	Data[111:96]	6
FRMUP[123:122]	EC[1:0]	EC
FRMUP[125:124]	IC[1:0]	
FRMUP[127:126]	H[1:0] = 2'b10	HFH[1:0] = 2'b10

7 groups of 4 input e-Ports
Number of data ports:

- 28 eLinks @ 160 Mbps
- 14 eLinks @ 320 Mbps
- 7 eLinks @ 640 Mbps

Note: This is how you will see the uplink frame after it has been decoded by the lpGBT-FPGA receiver. Not how it is actually transmitted by the lpGBT over the fiber.

Backend Links



Backend Links – TTC-PON

Based on ITU XG-PON Fiber-to-the-Home technology

- Upstream/downstream wavelength multiplexed
- Commercial components
- Custom protocol tailored to our specifications

Point-to-MultiPoint

- Split-ratio - 1:64

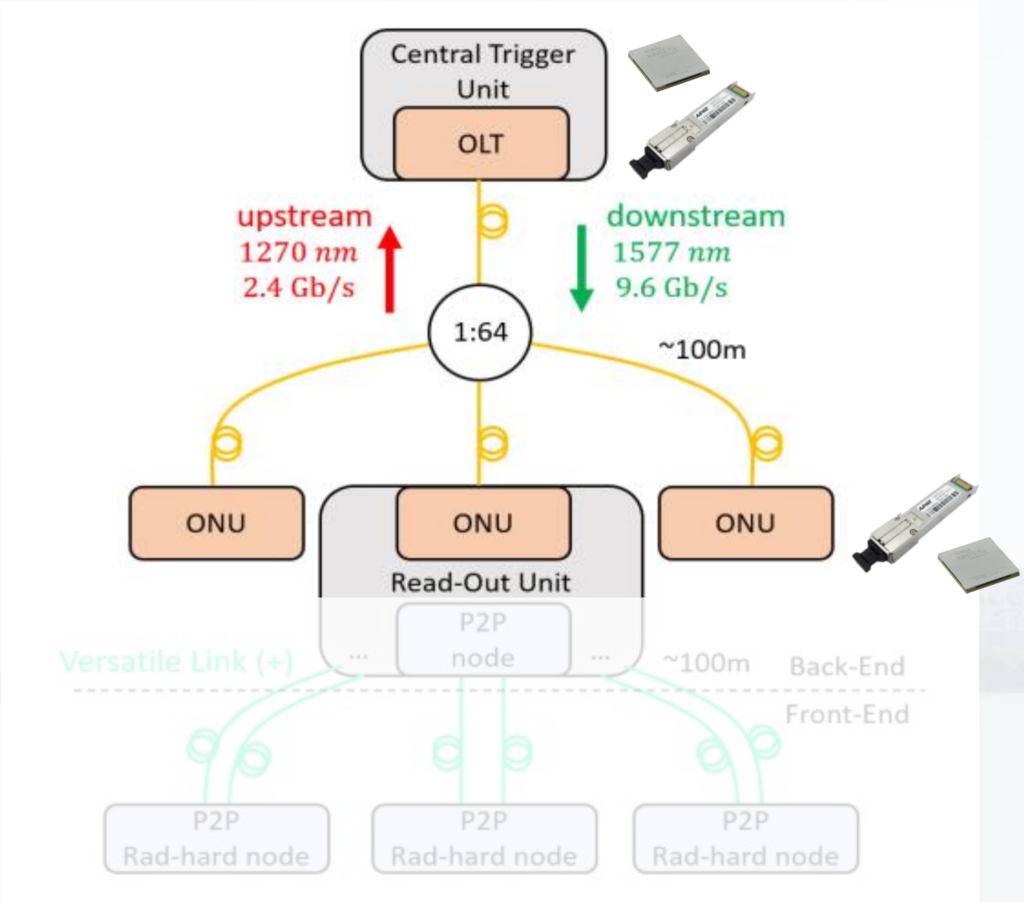
Payload (including FEC)

- Downstream Data-rate=9.6Gb/s
- Upstream Data-rate=2.4Gbps shared between ONUs

ONU recovered clock jitter - ~ 5ps rms

Latency (active components)

- Fixed and deterministic
- Downstream - ~100ns
- Upstream - ~ Number_ONU x 125ns



https://gitlab.cern.ch/TTC-PON/official_release

Backend Links – Point-to-Point

Timing

- 8b10b + timing optimizations

DAQ

- Aurora, Interlaken, custom protocols…

Based on COTS

- No multiplexing
- Commercial components
- Custom protocol tailored to our specifications

Data rates

- 10-56 Gbps

Recovered clock jitter: O(10) ps

- When proper PLLs and transceivers are used

Fixed and Deterministic Latency

- O(10) ps precision when TCLink IP is used on bidirectional links

High Performance Fixed latency Jitter Attenuators

Mid/High range FPGAs + TCLink IP

MM Minipods 12x10-25Gbps, 100m

MM Firefly 12x10-25Gbps, 100m

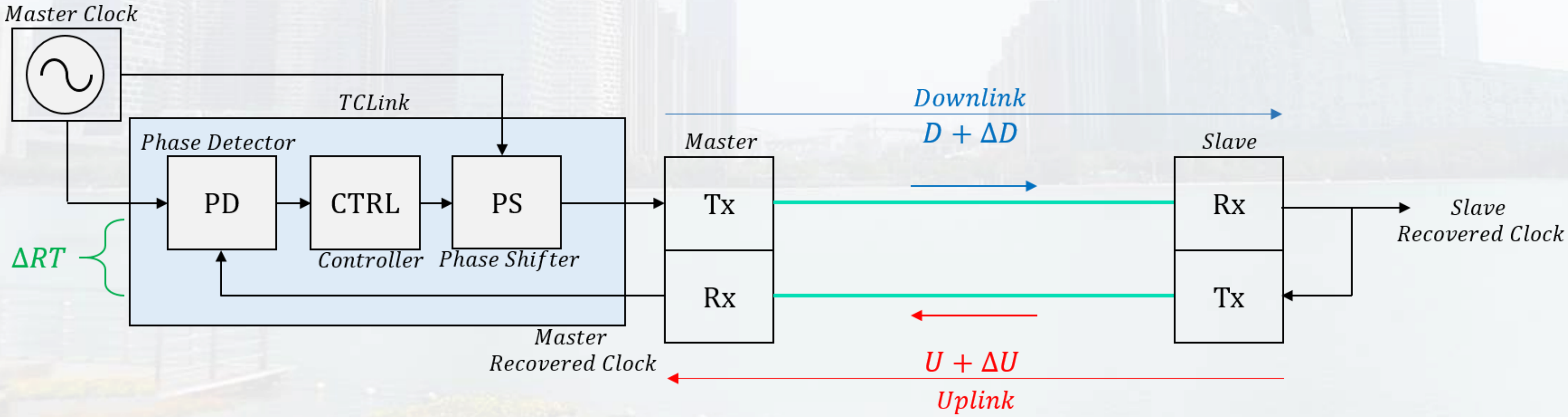
MM/SM SFP28 1x25Gbps, 100m-20km

MM/SM SFP28 4x25Gbps, 100m-20km

TCLink

https://gitlab.cern.ch/HPTD/tclink

- The TCLink is a protocol-agnostic FPGA core designed to mitigate long-term phase variations in high-speed optical links
  - Protocol-agnostic but already integrated with lpGBT
  - Specific to Ultrascale+ FPGAs

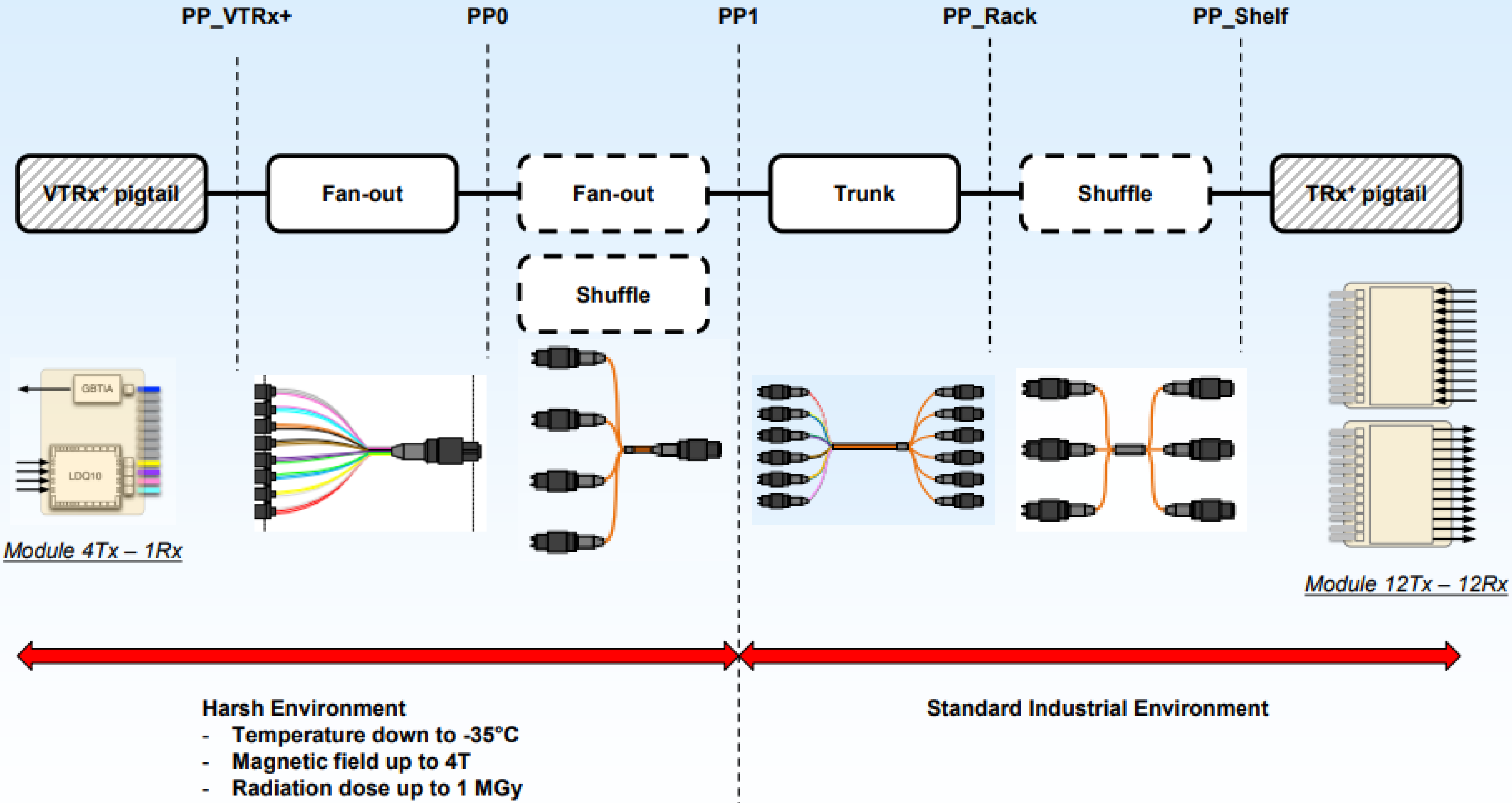


Outline

- Fundamental concepts
- Physical devices
- Applications
  - Frontend optical links
  - Backend optical links
- **Installation / Commissioning**

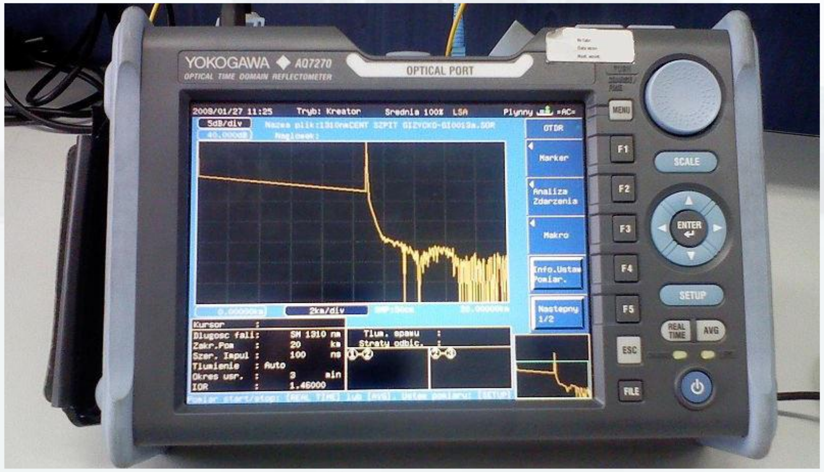
Cabling Plant

- Up to 150m, up to 5 breakpoints (VL+)
- The shorter the length and the fewer the breakpoints, the better the power margin

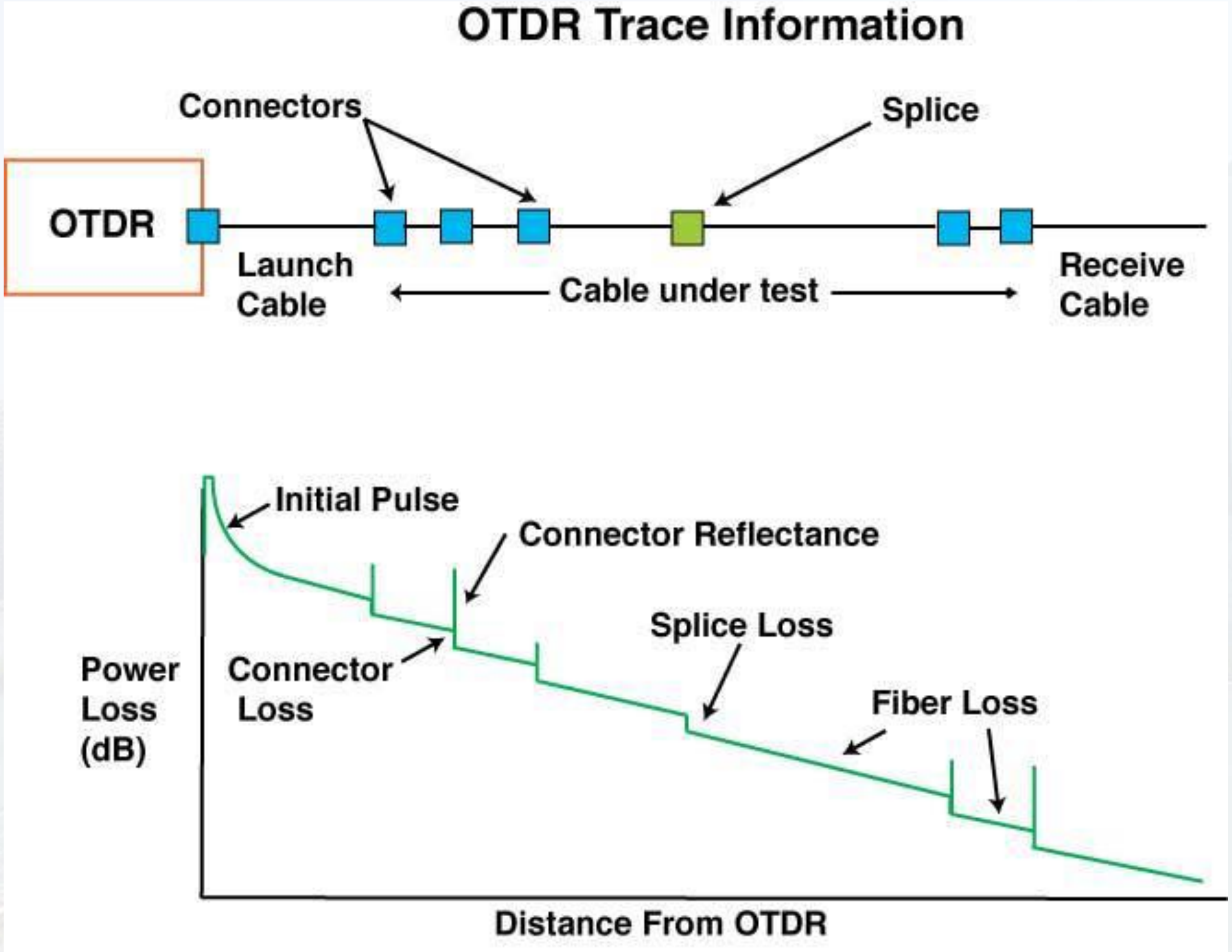


Optical Time-Domain Reflectometer (ODTR)

- Indispensable tool to test fiber installations
- Inject a train of light pulses in fiber strand
- Analyze light that is reflected back
- Characterize fiber path
  - Locations of fiber splices
  - Locations of fiber breaks
  - (As distance from probe)
  - Overall attenuation of fiber



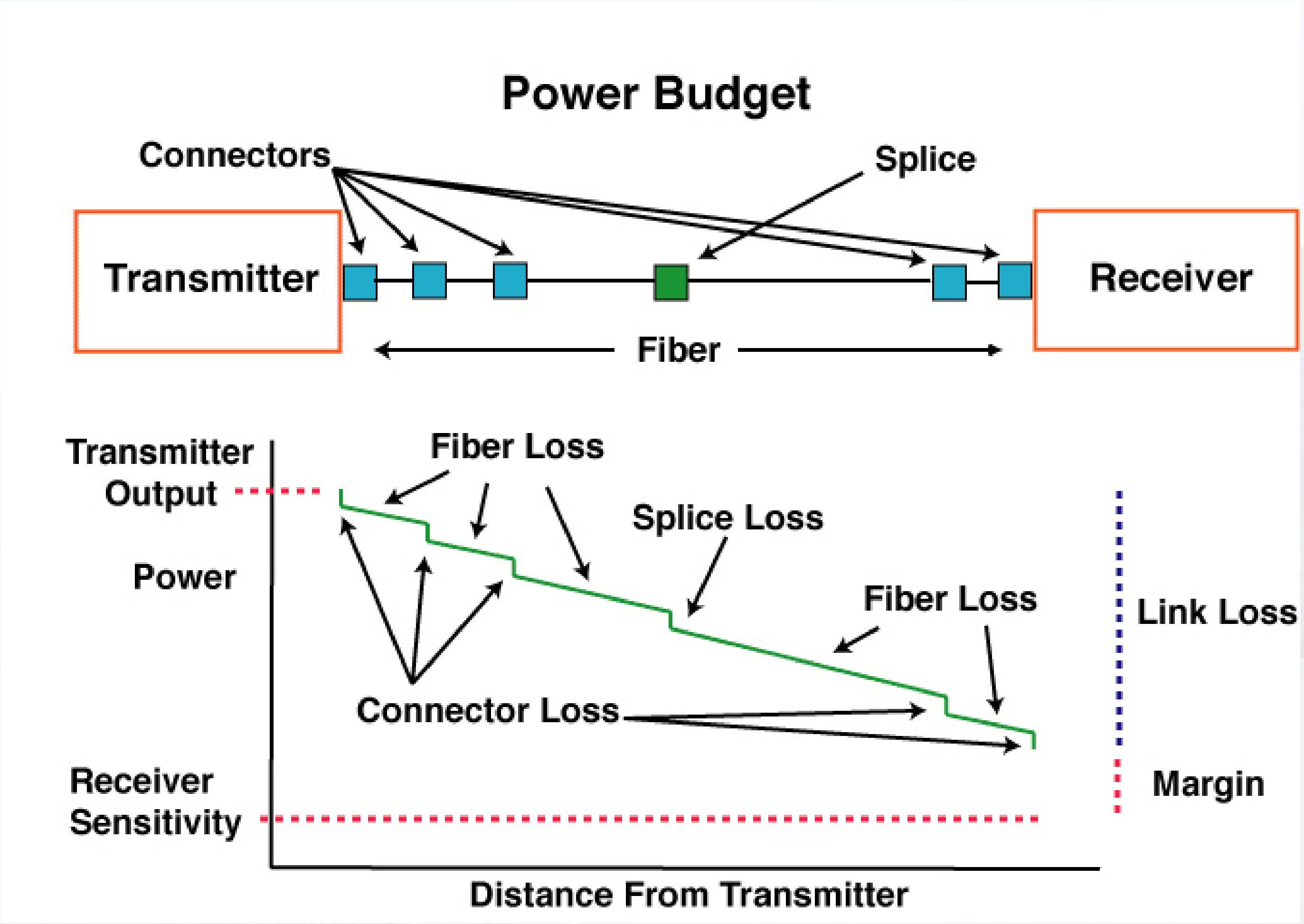
ODTR output diagram



Optical Power Meter (Light meter)

- Actually two devices
  - Optical reference
  - Receiver module
- Measures the “brightness” of the signal [dBm] or [mW]
- Automated (re)-calibration
  - Measure either absolute power or relative loss
- MPO polarity detection
- Scan one entire MPO bundle at once

Plan your optical budget…



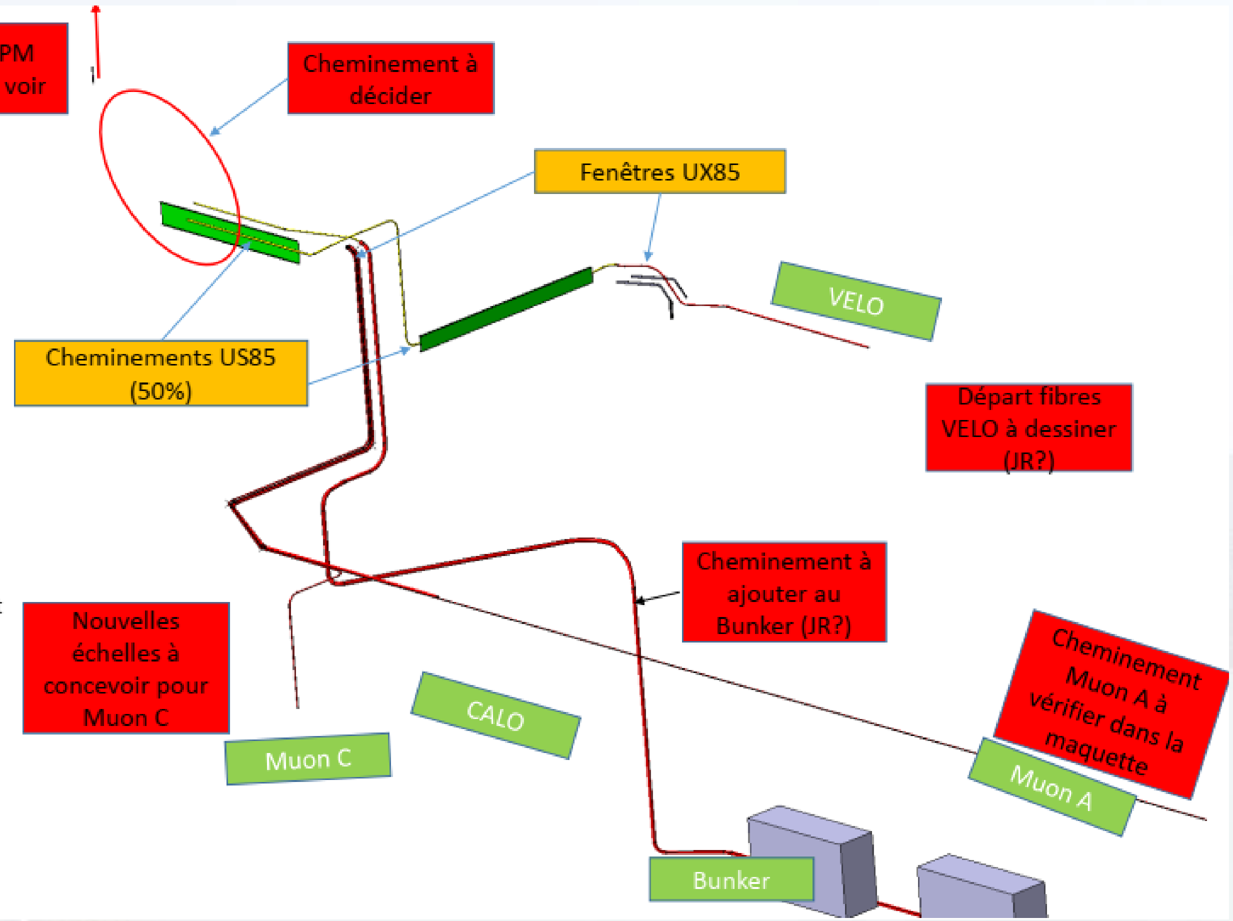
…plan your cable paths…

Fait:

- Partes pour chaque partie du cheminement créées
- Cheminement décidé à 90%
- Design nouvelles échelles US 50%

A faire:

- Concevoir cheminement du US au PM
- Finir échelles US, UX, Bunker
- Fixation dans PM à décider
- Ajouter détails – épaisseurs et nombres de fibres
- <u>Calculer longueur de chaque fibre</u>



…plan your cabling and spares…

Custom SQL query returning 101 rows (hide)

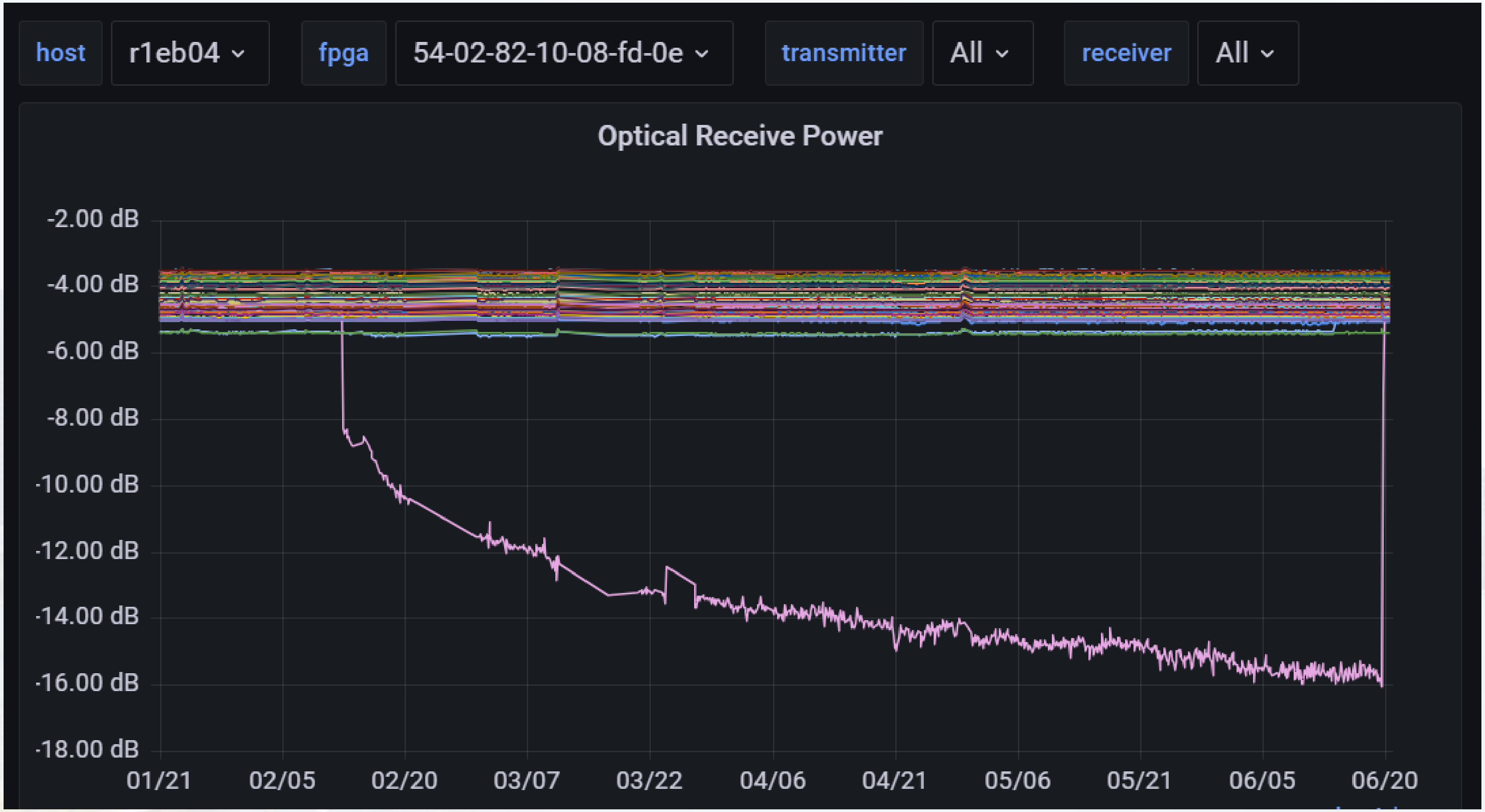
```
select trunk_name, trunk_fiber, trunk_mpo_port, trunk_mpo_fiber, pcie40_name, pcie40_mpo_port,
pcie40_mpo_fiber, pcie40_link_dir, pcie40_link_number, sd_name, notes from fibers order by trunk_name,
trunk_fiber limit 101
```

Format SQL | Run SQL

This data as json, CSV

trunk_name	trunk_fiber	trunk_mpo_port	trunk_mpo_fiber	pcie40_name	pcie40_mpo_port	pcie40_mpo_fiber	pcie40_link_dir	pcie40_link_number	sd_nar
220-01	1	1	1	CASOL011	1	1	Rx	0	EC00_
220-01	2	1	2	CASOL011	2	1	Tx	0	EC00_
220-01	3	1	3	CASOL011	1	2	Rx	1	EC01_
220-01	4	1	4	CASOL011	2	2	Tx	1	EC01_
220-01	5	1	5	CASOL011	1	3	Rx	2	EC02_
220-01	6	1	6	CASOL011	2	3	Tx	2	EC02_
220-01	7	1	7	CASOL011	1	4	Rx	3	EC03_
220-01	8	1	8	CASOL011	2	4	Tx	3	EC03_
220-01	9	1	9	CASOL011	1	5	Rx	4	EC04_
220-01	10	1	10	CASOL011	2	5	Tx	4	EC04_

...and monitor your links



Optimizing your TDAQ infrastructure

- Constraints: material budget, power, cost…
- Organization: frontend vs. backend may not be the same people
  - An optical cabling strategy that is optimal for the frontend, might not be a globally optimal solution
  - Shifting some complexity into the backend, may result in a system that is, overall, less cost-effective
  - Avoid overoptimizing along one specific direction (link utilization, number of fibers) without measuring global ramifications
  - Simulate, simulate again, talk to your colleagues, add contingencies, plan for emergencies