Extending 100Gbit/s Ethernet

Ariën Vijn
arien.vijn@ams-ix.net

Agenda

- AMS-IX
- 100Gbit/s technology
- Problem statement
- Optical Amplifier development
- Metro DWDM equipment

Agenda

- AMS-IX
- AMS-IX
- 100Gbit/s technology
- Problem statement
- Optical Amplification
- Metro DWDM equipment

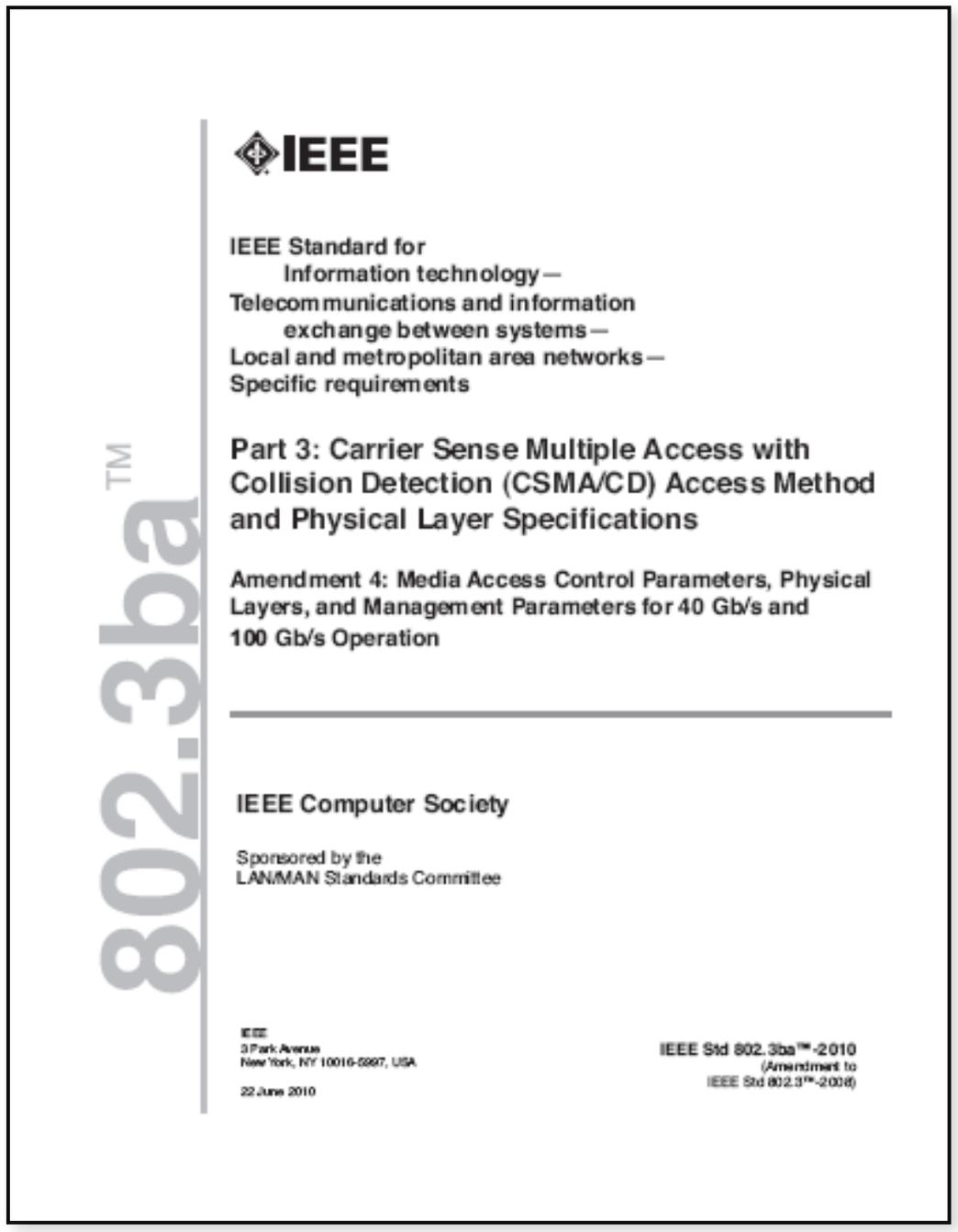
AMS-IX

- Amsterdam Internet Exchange
  - Not for profit organization.
  - 516 Networks (ASes).
  - 1857Gbit/s peak
  - 911Gbit/s average over the last 16 months
  - 11 Operational sites.
    - We will add one more site this year.
  - 24 x 100Gbit/s backbone links.

Agenda

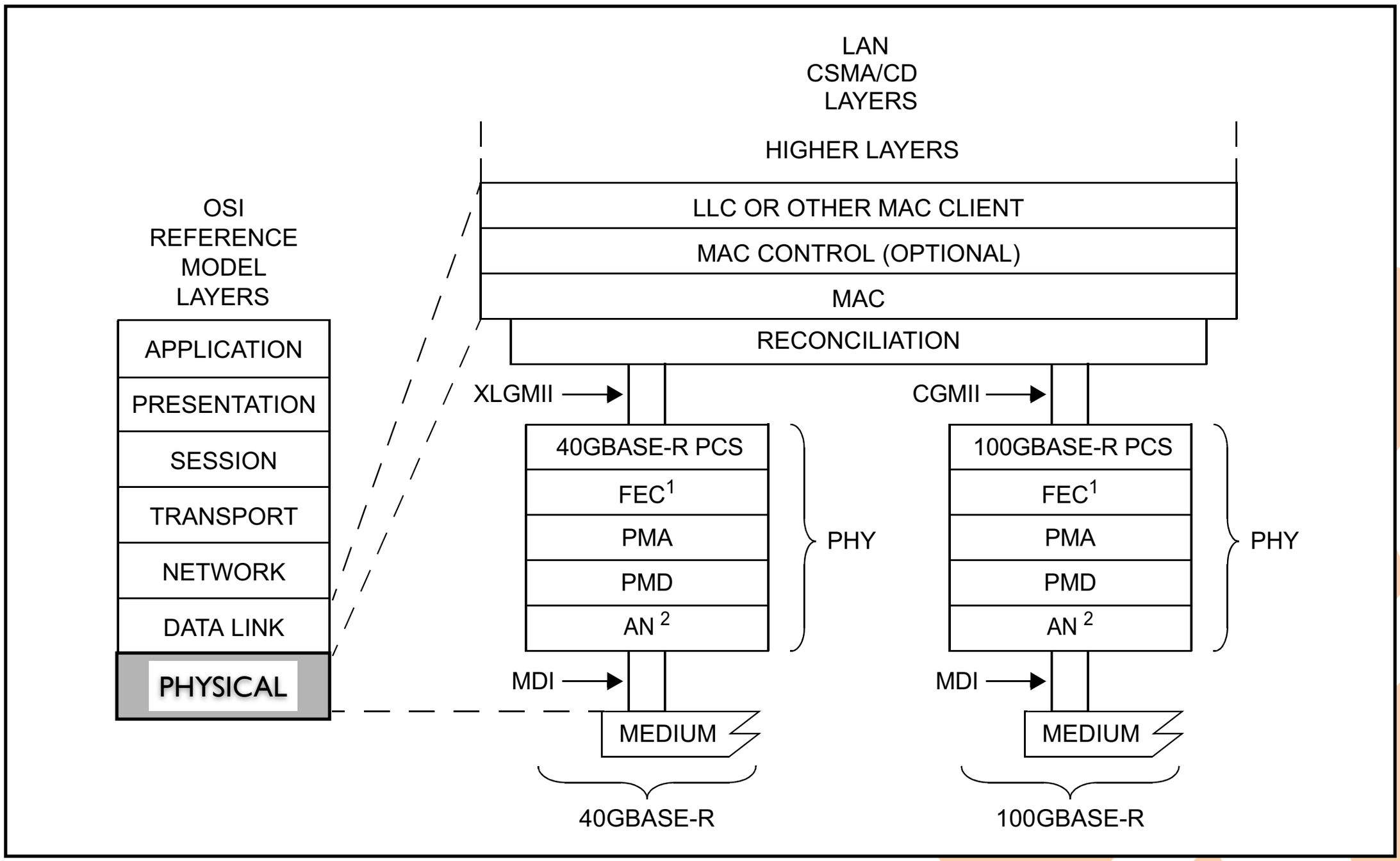
- AMS-IX
- 100Gbit/s technology
- Problem statement
- Optical Amplifier development
- Metro DWDM equipment

IEEE 802.3ba

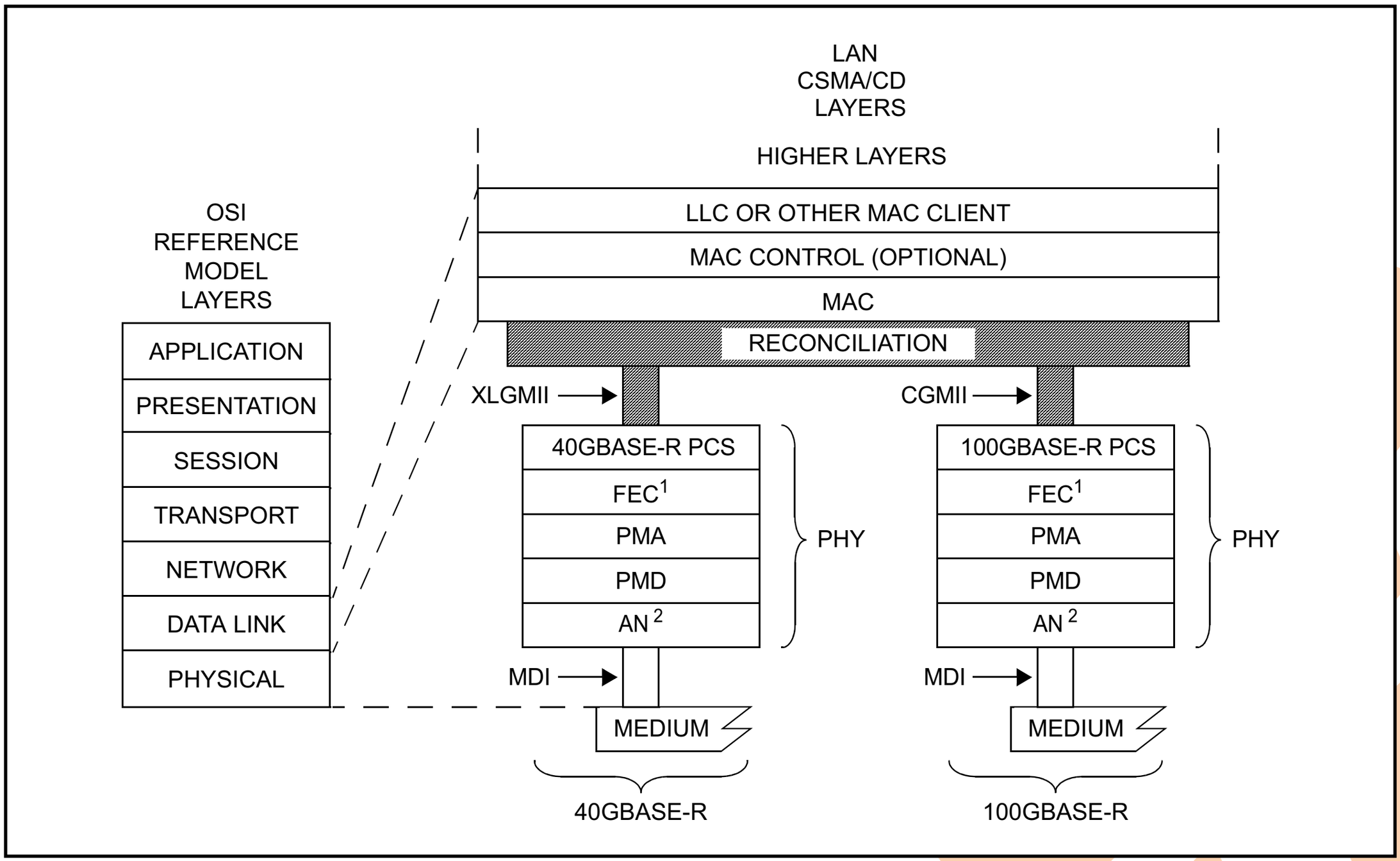


- June 17, 2010
  - IEEE 802.3ba was approved
  - 40GE / 100GE
- 100G
  - 100GBASE-SR10 (< 150m, MMF)
  - 100GBASE-LR4 (< 10km, SMF)
  - 100GBASE-ER4 (< 40km, SMF)

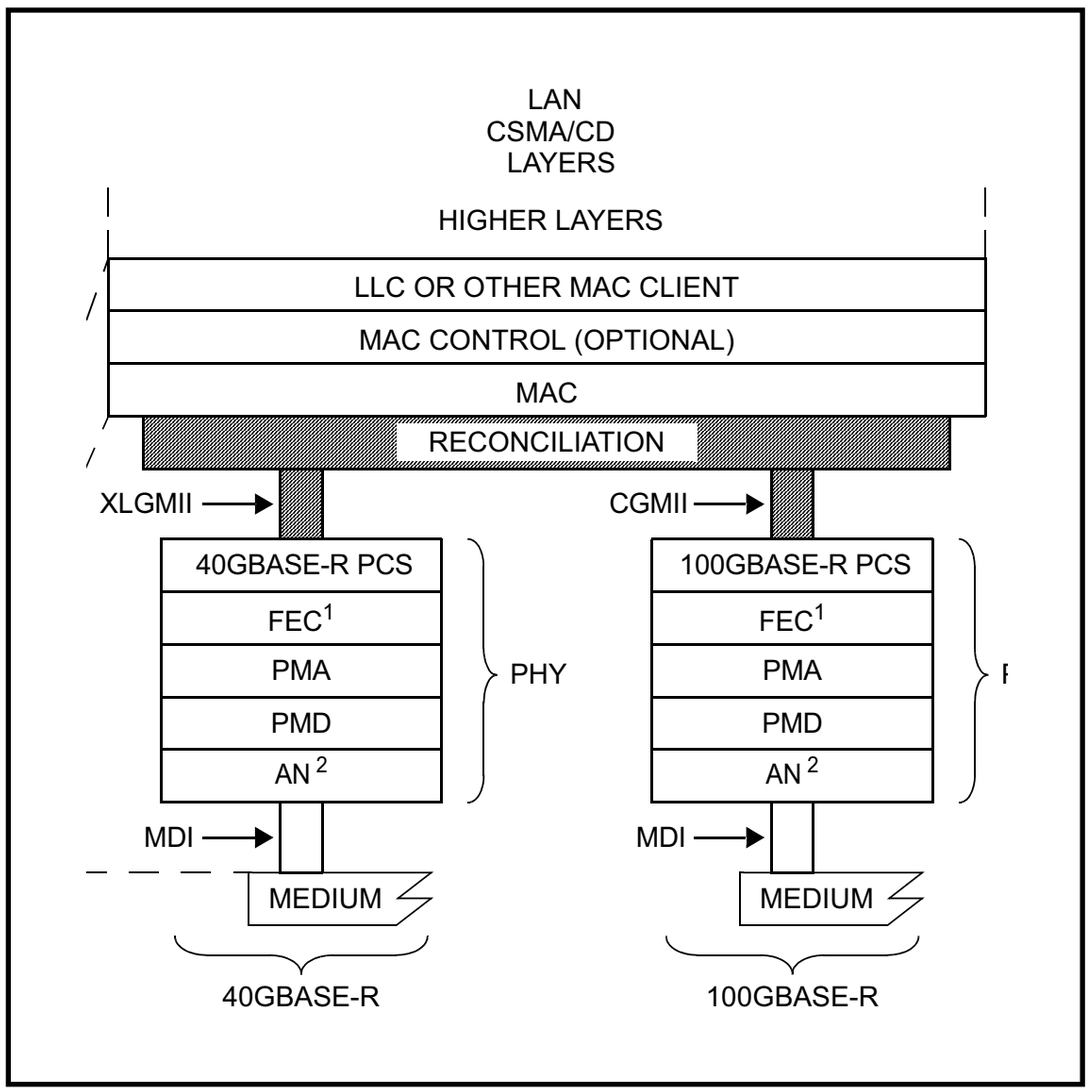
IEEE 802.3ba


LAN
CSMA/CD
LAYERS
HIGHER LAYERS
LLC OR OTHER MAC CLIENT
MAC CONTROL (OPTIONAL)
MAC
RECONCILIATION
OSI
REFERENCE
MODEL
LAYERS
APPLICATION
PRESENTATION
SESSION
TRANSPORT
NETWORK
DATA LINK
PHYSICAL
XLGMII
CGMII
40GBASE-R PCS
FEC[1]
PMA
PMD
AN [2]
PHY
100GBASE-R PCS
FEC[1]
PMA
PMD
AN [2]
PHY
MDI
MDI
MEDIUM
MEDIUM
40GBASE-R
100GBASE-R

IEEE 802.3ba

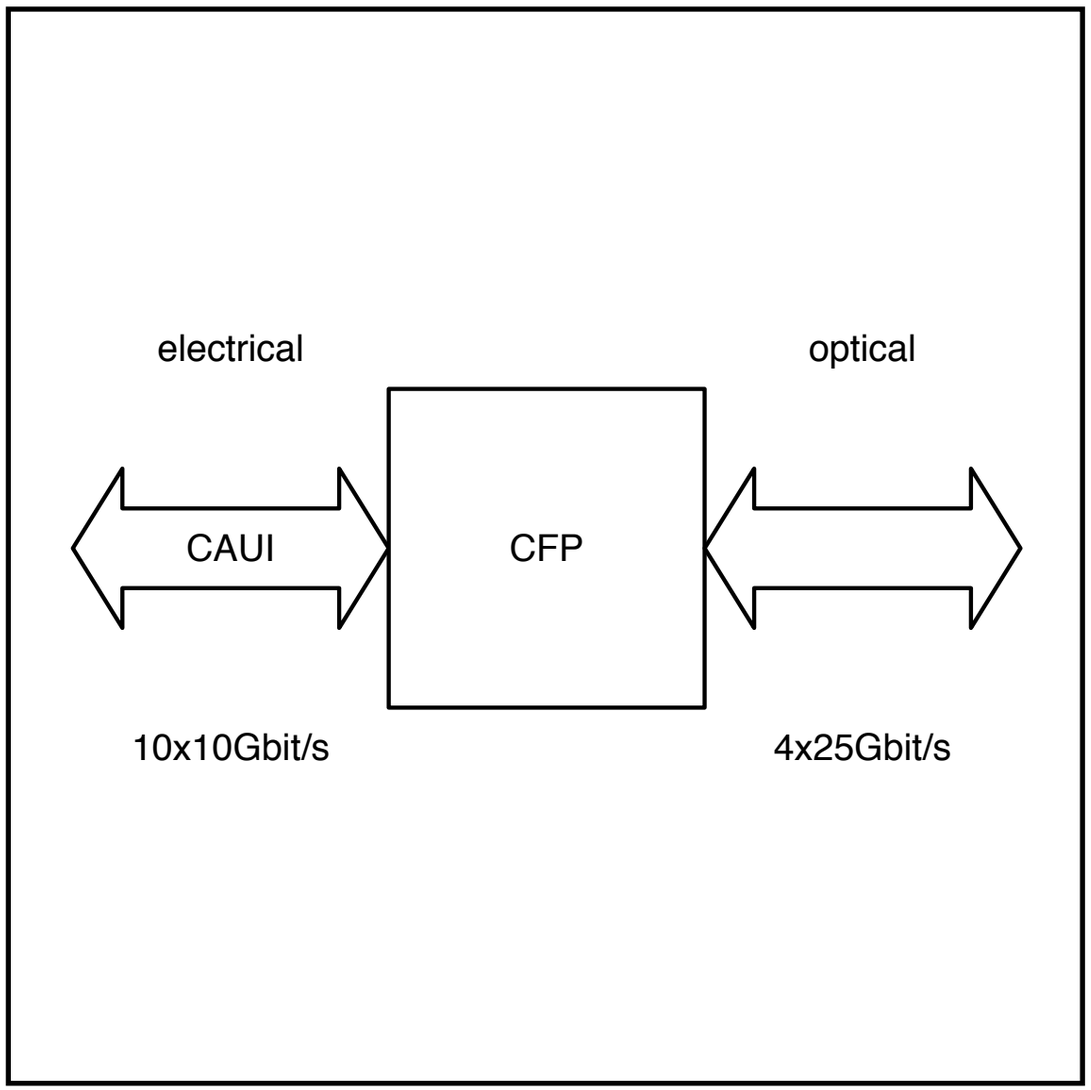

LAN
CSMA/CD
LAYERS
HIGHER LAYERS
LLC OR OTHER MAC CLIENT
MAC CONTROL (OPTIONAL)
MAC
RECONCILIATION
XLGMII
CGMII
40GBASE-R PCS
FEC[1]
PMA
PMD
AN[2]
PHY
100GBASE-R PCS
FEC[1]
PMA
PMD
AN[2]
PHY
MDI
MDI
MEDIUM
MEDIUM
40GBASE-R
100GBASE-R
OSI
REFERENCE
MODEL
LAYERS
APPLICATION
PRESENTATION
SESSION
TRANSPORT
NETWORK
DATA LINK
PHYSICAL

IEEE 802.3ba



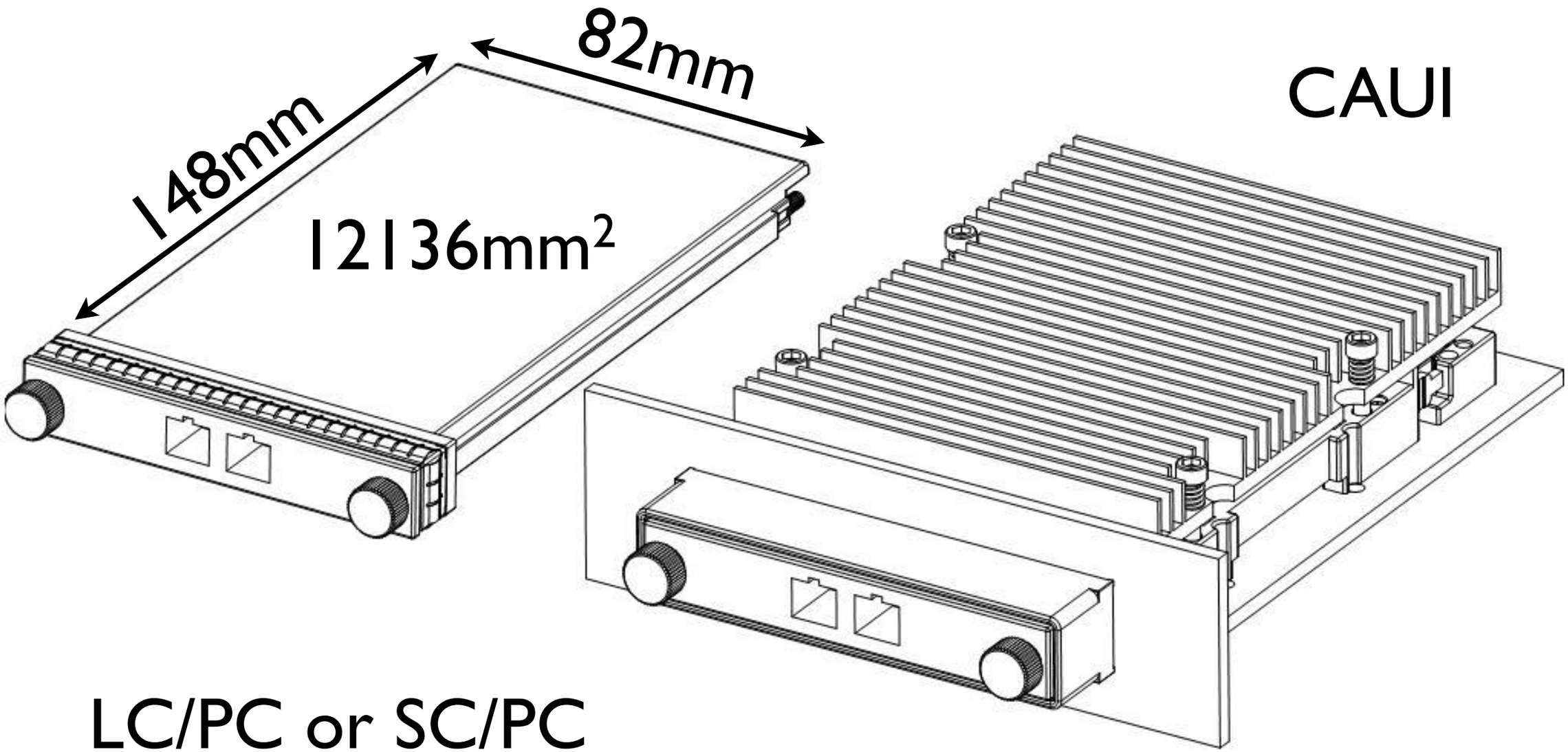
- 100Gbit/ Electrical interface
  - CAUI (CFP)
    - 10 x 10Gbit/s
  - CAUI-4 (CFP-2)
    - 4 x 25Gbit/s
  - cPPI-4 (CFP-4)
    - 4 x 25Gbit/s

CFP

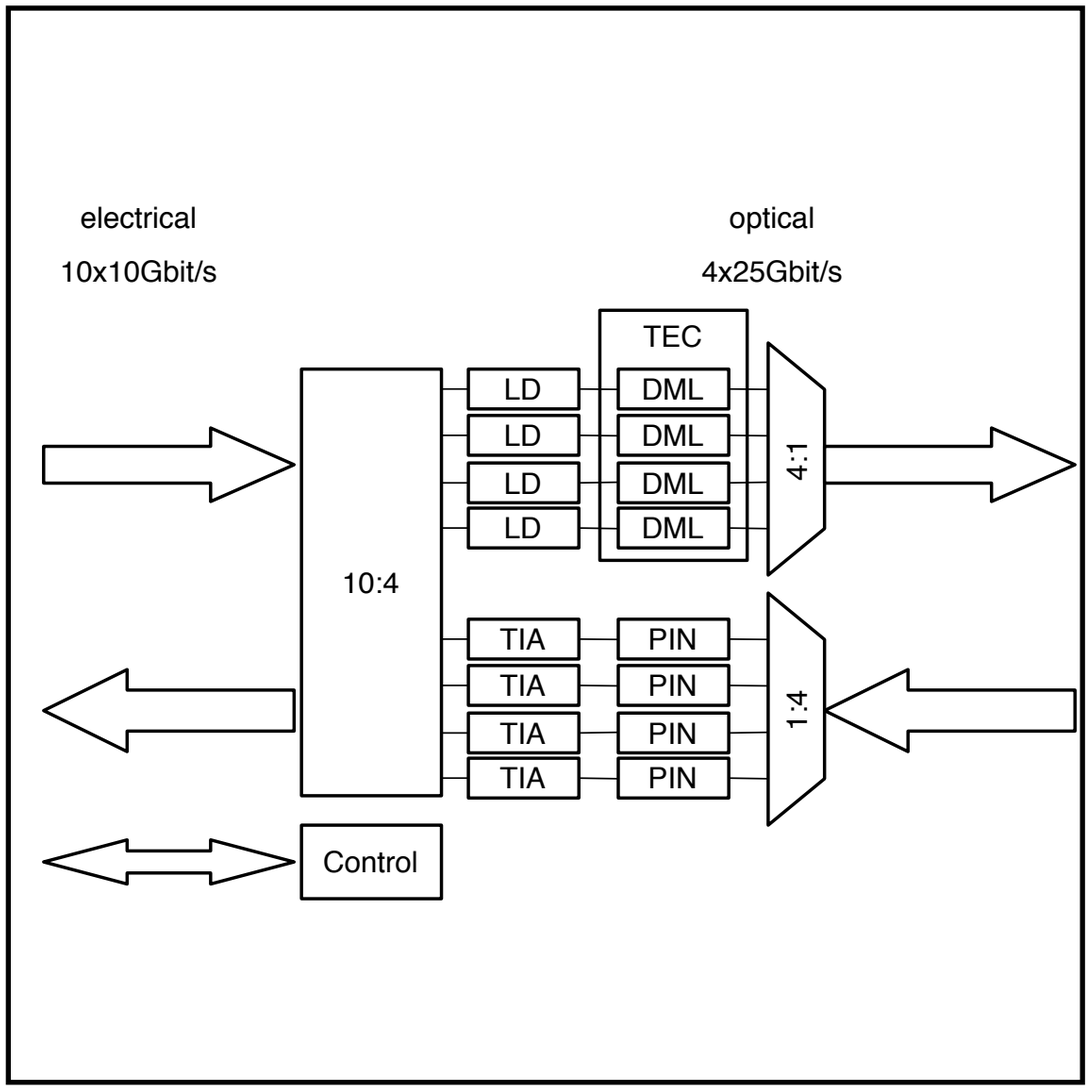


- C Form-factor Pluggable (CFP)
- 100Gbit/ Electrical interface
  - CAUI
    - CFP
    - 10 x 10Gbit/s
- Optical Interface
  - Multi mode 10x10Gbit/s
  - Single mode 4x25Gbit/s

CFP



CFP

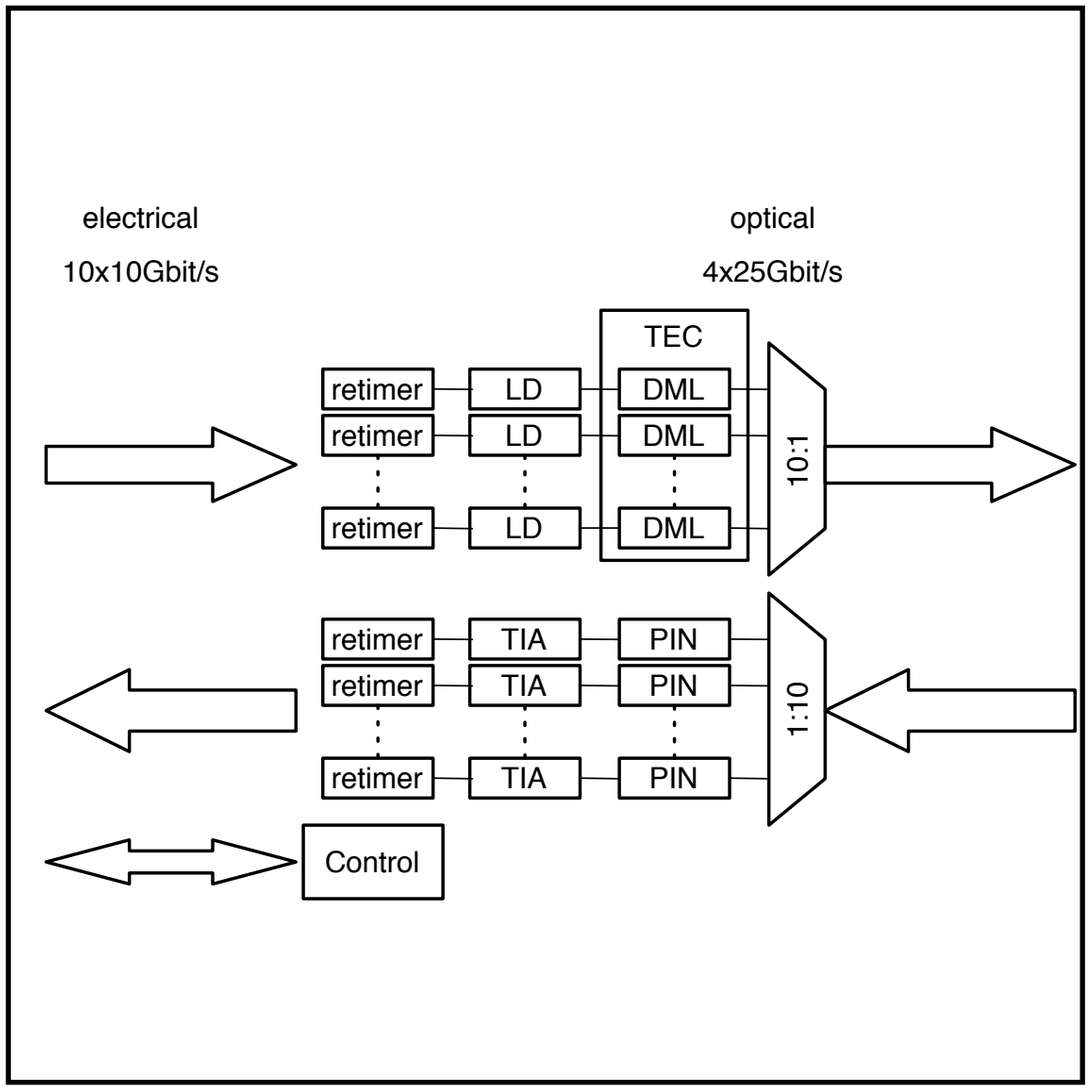


- 10:4 / 4:10 electrical “gearbox”
- 4 transmitters
- 4 receivers
- 4:1 / 1:4 optical muxes
- Control unit
- This is a lot of elements in one transceiver

10x10 MSA

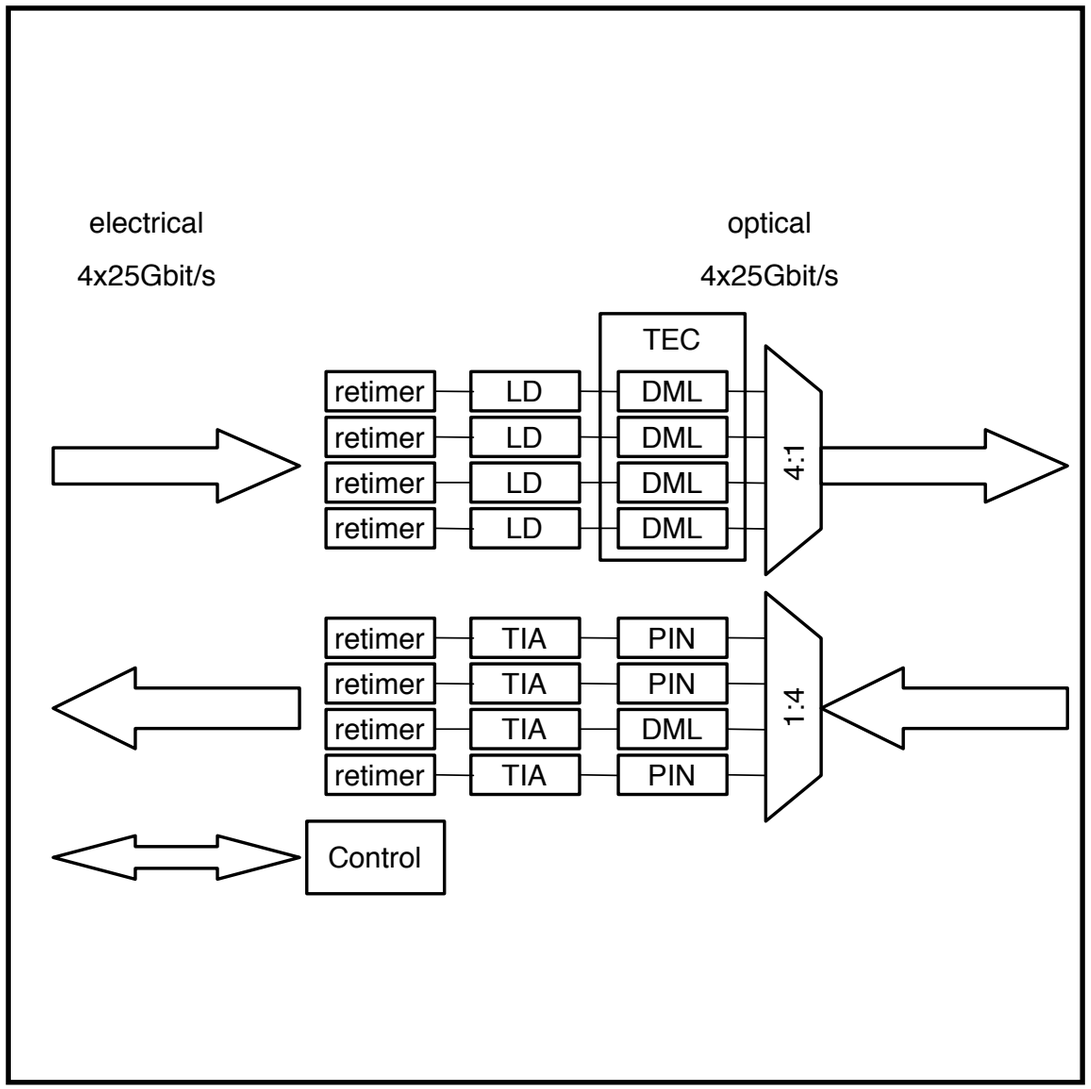
- Outside IEEE initiative
- 26 participants
  - Network operators
    - AMS-IX among others
  - No Cisco, Juniper, Alcatel
    - Works in Juniper though
- Less expensive 100GBASE-LR4 between:
  - 100GBASE-SR10 supports up to 150m (OM4 MMF)
  - 100GBASE-LR4 supports up to 10km (SMF)

10x10 MSA



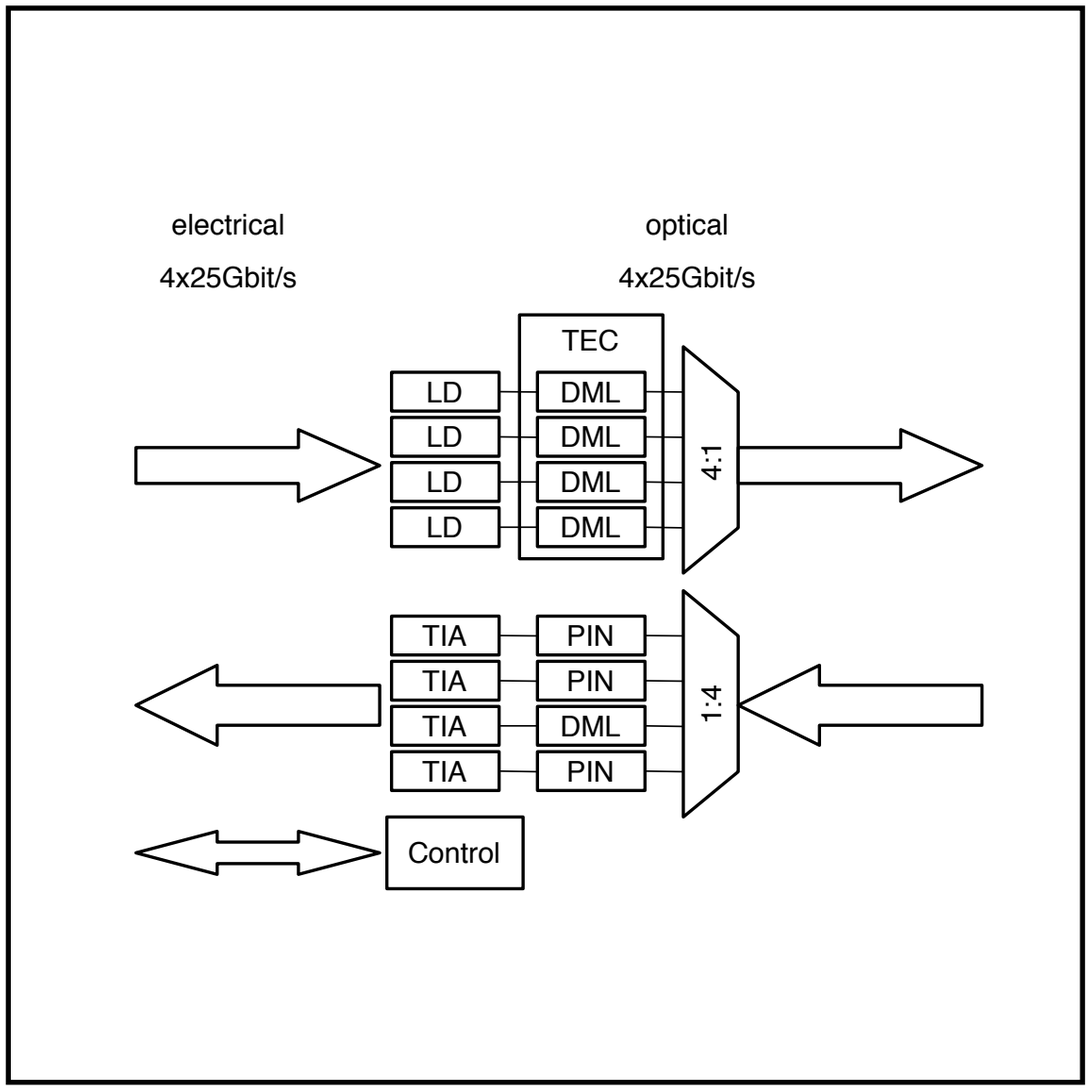
- 10x10Gbit/s electrical and optical
  - 10 Lasers
  - 10 Transceivers
  - 10:1 / 1:10 optical muxes
- No “gearbox”
  - Retiming circuits instead
- Considerably cheaper LR-4
- Less power hungry than LR-4

CFP-2 (future)



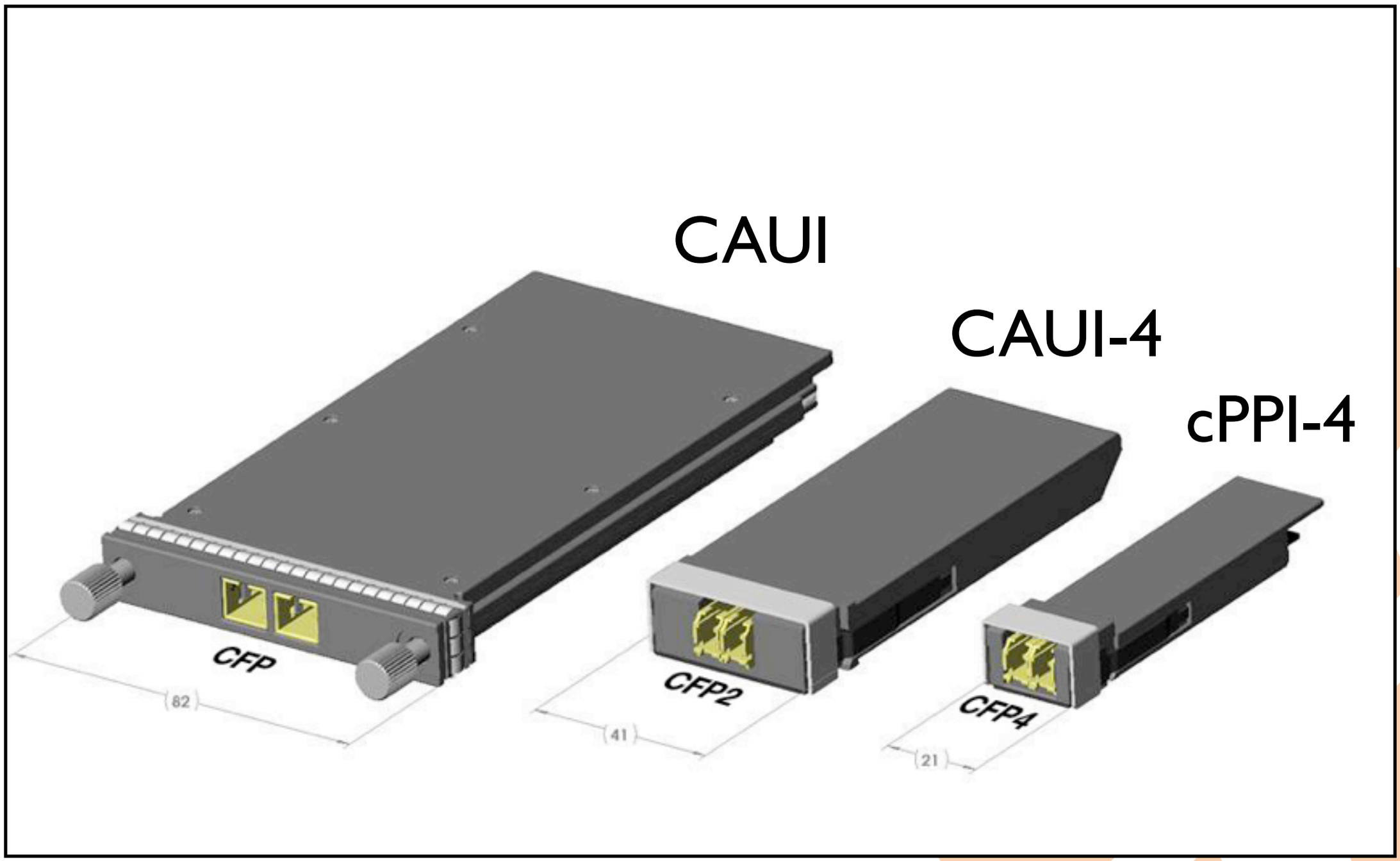
- Smaller modules than CFPs
- 4x25Gbit/s electrical and optical
  - 4 Lasers
  - 4 Transceivers
  - 4:1 / 4:1 optical muxes
- No “gearbox”
- Still Retiming circuits
- Cheaper than CFPs
- Less power hungry than CFPs
- 10x10 MSA would require a gearbox

CFP-4 (future)



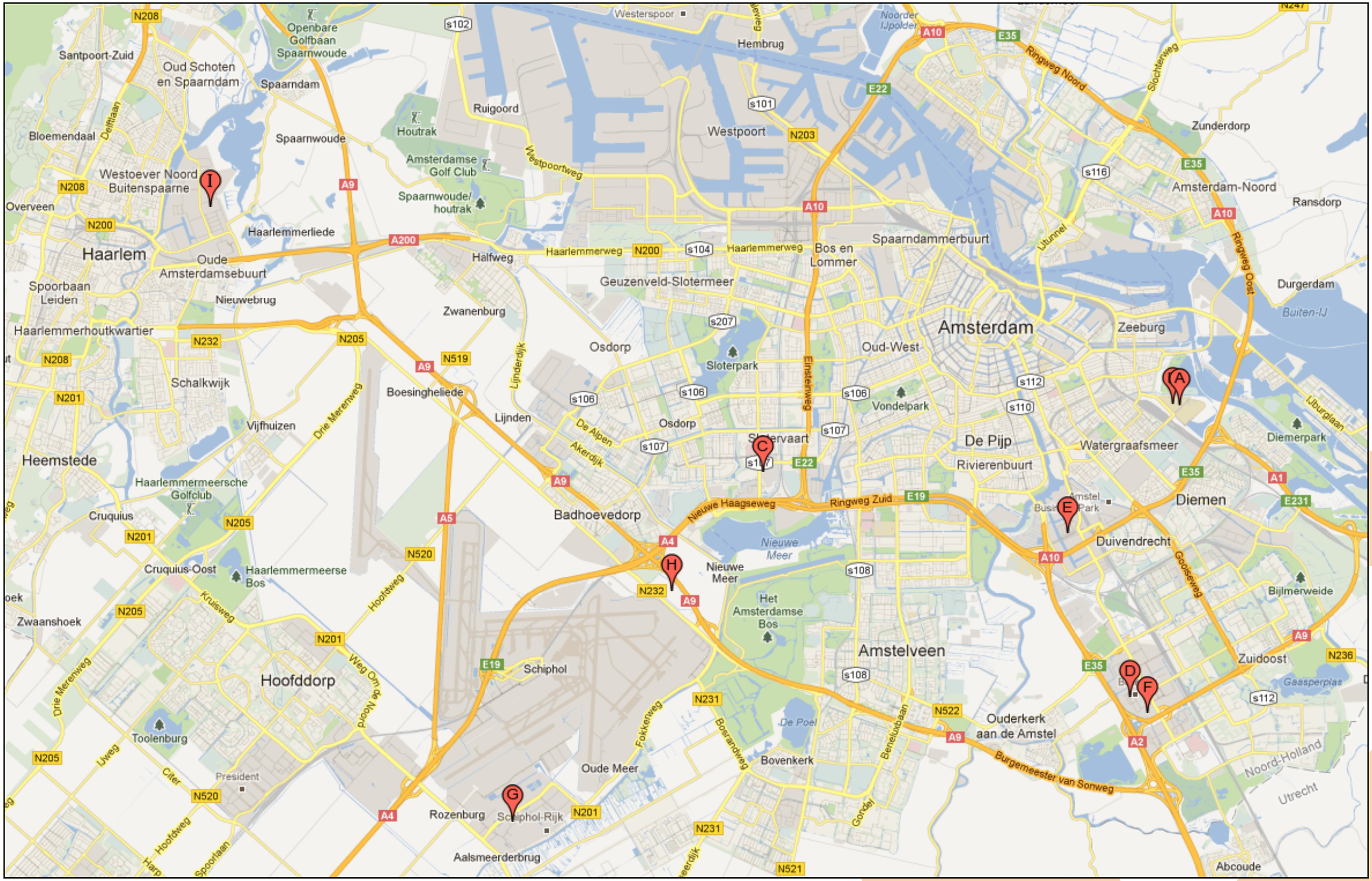
- Smaller modules than CFP-2
- 4x25Gbit/s electrical and optical
  - 4 Lasers
  - 4 Transceivers
  - 4:1 / 4:1 optical muxes
- No retiming units
- Require external retiming units.

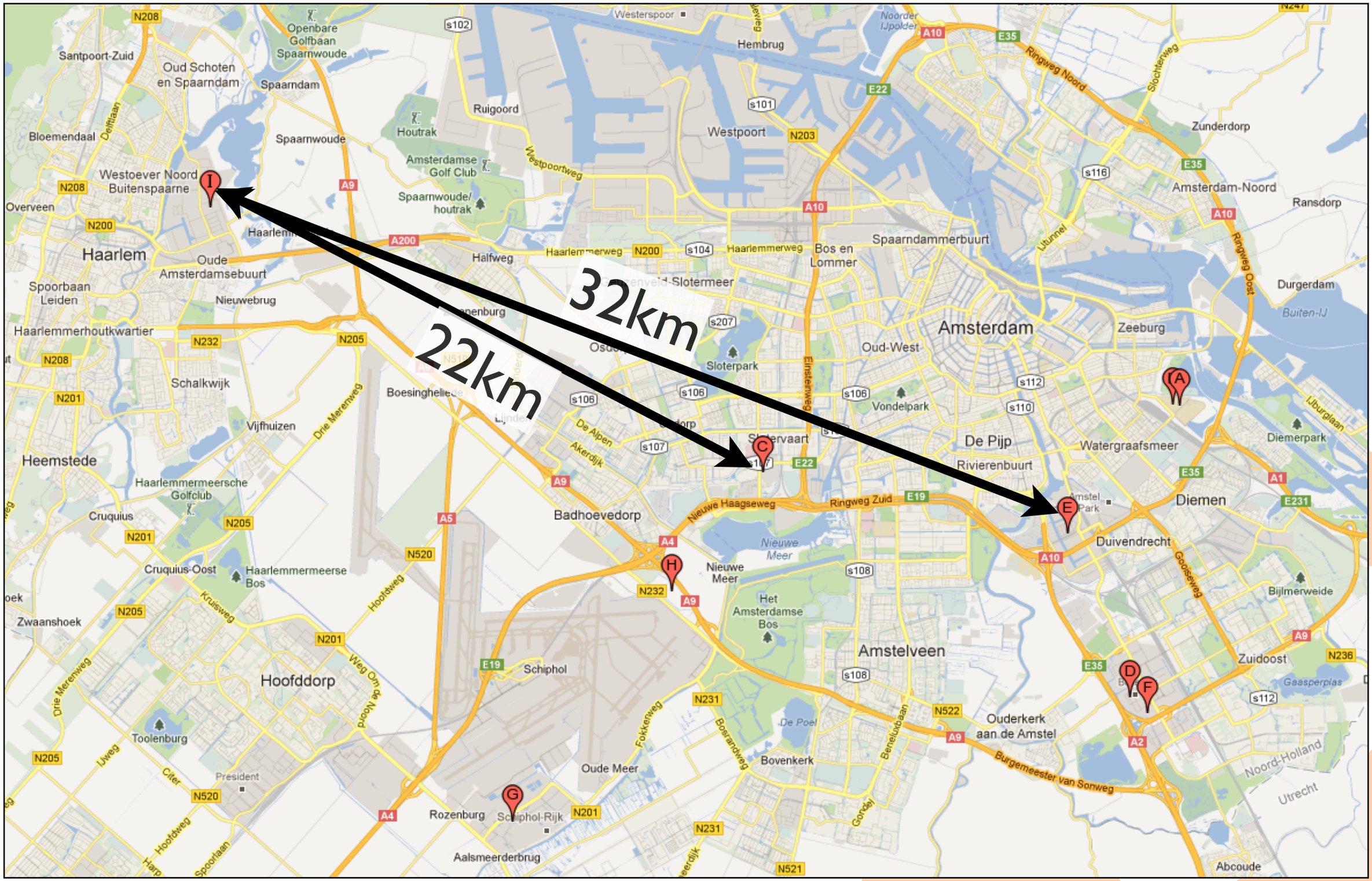
CFP2, CFP4



Agenda

- AMS-IX
- 100Gbit/s technology
- Problem statement
- Optical Amplifier development
- Metro DWDM equipment


Westerspoor
Hembrug
Noorder IJpolder
A10
E35
Ringweg Noord
Slochterweg
E22
s101
Westpoort
N203
Zunderdorp
Openbare Golfbaan Spaarnwoude
Santpoort-Zuid
Oud Schoten en Spaarndam
Spaarndam
Ruigoord
Houtrak
Bloemendaal
Delftlaan
Spaarnwoude
Amsterdamse Golf Club
Westpoortweg
Westoever Noord Buitenspaarne
N208
A9
Spaarnwoude/ houtrak
s116
Amsterdam-Noord
Ransdorp
Overveen
N200
A10
Haarlemmerliede
A200
Haarlem
Haarlemmerweg
N200
s104
Haarlemmerweg
Bos en Lommer
Spaarndammerbuurt
Halfweg
Oude Amsterdamsebuurt
Geuzenveld-Slotermeer
Ringweg Oost
Durgerdam
Spoorbaan Leiden
Nieuwebrug
Zwanenburg
Buiten-IJ
Haarlemmerhoutkwartier
s207
Amsterdam
Zeeburg
N232
N205
Osdorp
Oud-West
N208
A9
N519
Sloterpark
s112
Schalkwijk
Boesinghelaan
s106
s106
s106
Vondelpark
s110
N201
Lijnderdijk
Lijnden
Vijfhuizen
Drie Merenweg
De Alpen
Osdorp
Ijburglaan
Diemerpark
s107
De Pijp
Watergraafsmeer
Heemstede
Akerdijk
s107
Slotervaart
E22
Rivierenbuurt
E35
A1
Haarlemmermeersche Golfclub
A9
Ringweg Zuid
E19
Amstel Park
Diemen
Nieuwe Haagseweg
E231
Cruquius
N205
Badhoevedorp
A5
N201
A4
Duivendrecht
Nieuwe Meer
Cruquius-Oost
Haarlemmermeerse Bos
N520
Nieuwe Meer
s108
A10
Goosewg
Bijlmerweide
Kruisweg
Hoofdweg
Zwaanshoek
N205
Het Amsterdamse Bos
A9
N232
N201
Amstelveen
Zuidoost
N236
Hoofddorp
E19
Schiphol
E35
Gaasperplas
Weg Om de Noord
Drie Merenweg
N231
s112
Fokkerweg
Bosrandweg
De Poel
N522
Ouderkerk aan de Amstel
Toolenburg
Beneluxbaan
A9
A2
N205
Uweg
Oude Meer
Bovenkerk
Noord-Holland
Citer
President
N520
Burgemeester van Sonweg
Utrecht
Hoofdweg
A4
Rozenburg
Schiphol-Rijk
N201
Gondel
N231
Spoorlaan
Aalsmeerderbrug
N521
Abcoude
I
A
C
E
H
D
F
G

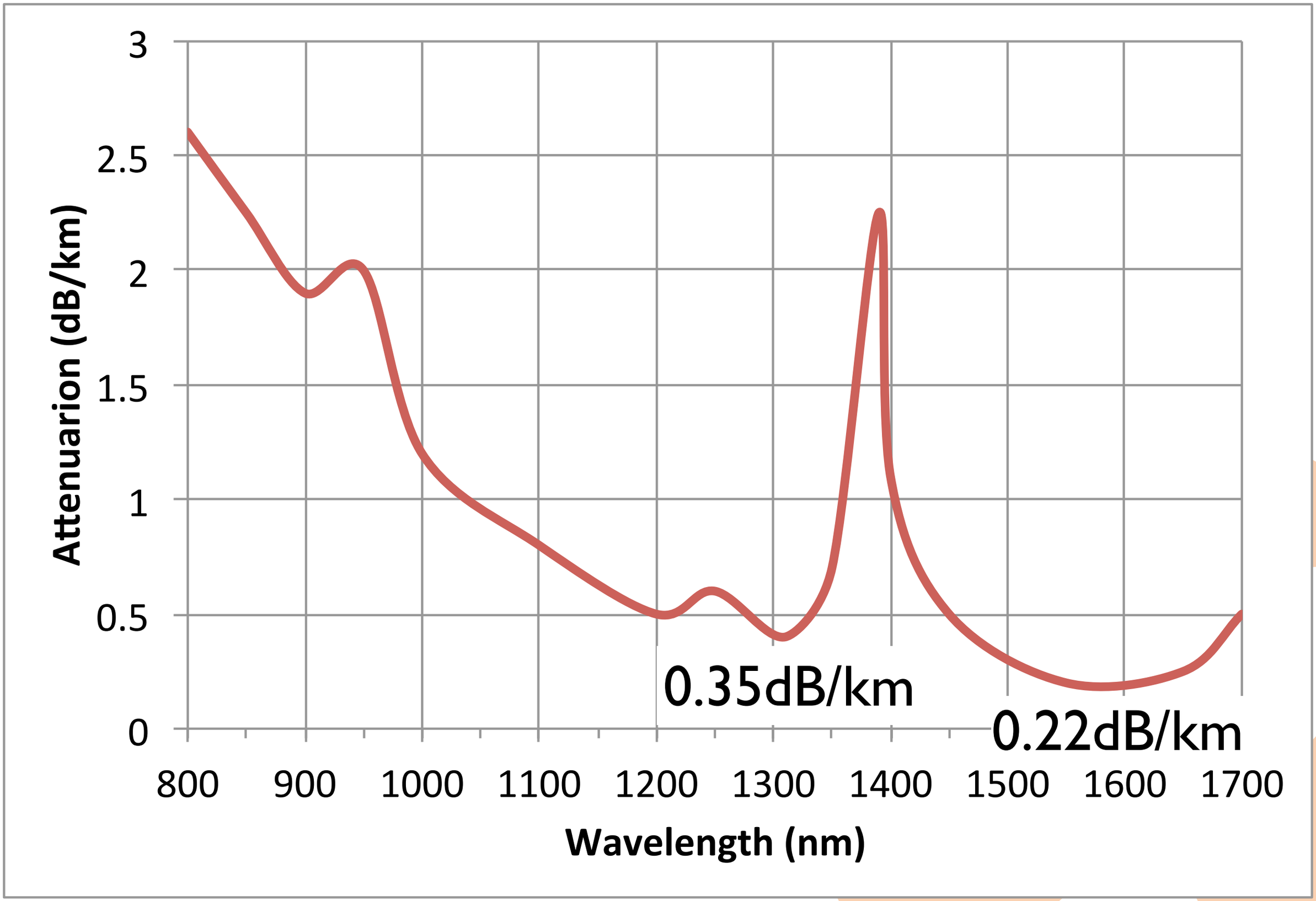

32km
22km
I
C
E
A
H
G
D
F
Westerspoor
Hembrug
Noorder IJpolder
Santpoort-Zuid
Oud Schoten en Spaarndam
Openbare Golfbaan Spaarnwoude
Spaarndam
Ruigoord
Houtrak
Westpoort
Zunderdorp
Bloemendaal
Spaarnwoude
Amsterdamse Golf Club
Westoever Noord Buitenspaarne
Spaarnwoude/ houtrak
Amsterdam-Noord
Ransdorp
Overveen
Haarlem
Oude Amsterdamsebuurt
Halfweg
Spaarndammerbuurt
Bos en Lommer
Durgerdam
Spoorbaan Leiden
Nieuwebrug
Haarlemmerhoutkwartier
Amsterdam
Zeeburg
Oud-West
Sloterpark
Schalkwijk
Vondelpark
Vijfhuizen
Boesinghelied
De Pijp
Watergraafsmeer
Diemerpark
Heemstede
Rivierenbuurt
Haarlemmermeersche Golfclub
Cruquius
Badhoevedorp
Diemen
Duivendrecht
Nieuwe Meer
Cruquius-Oost
Haarlemmermeerse Bos
Bijlmerweide
Zwaanshoek
Het Amsterdamse Bos
Amstelveen
Zuidoost
Hoofddorp
Schiphol
Toolenburg
De Poel
Ouderkerk aan de Amstel
Oude Meer
Bovenkerk
Noord-Holland
Rozenburg
Schiphol-Rijk
Utrecht
Aalsmeerderbrug
Abcoude

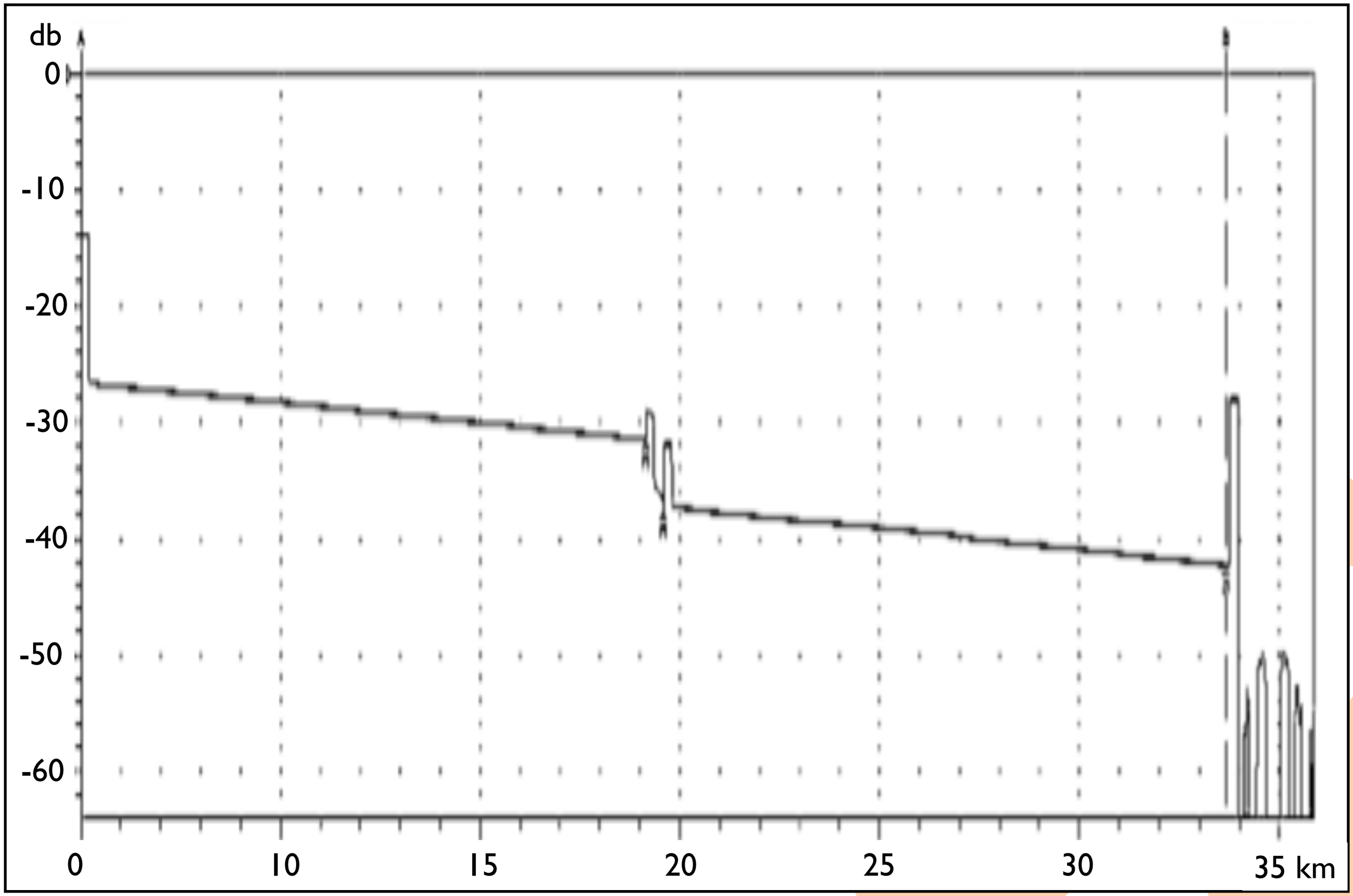
Problem Statement

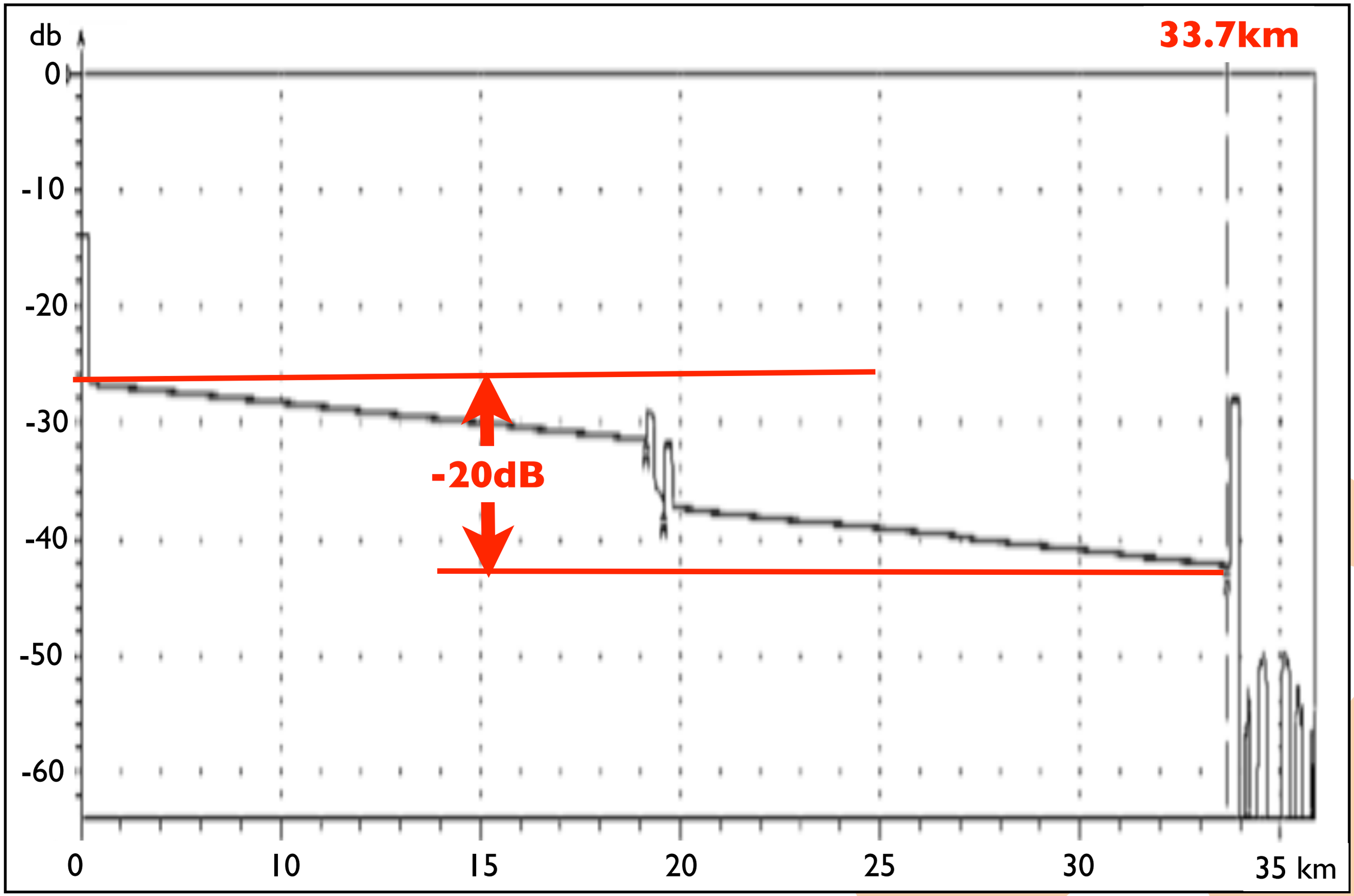
- Current 100Gb/s CFPs:
  - 10x10 for 2km.
  - 100G-LR4 for 10 km.
- Many of our links are too long for the available 100Gbit/s optics.
  - We need to drive up to 32km of fibre:

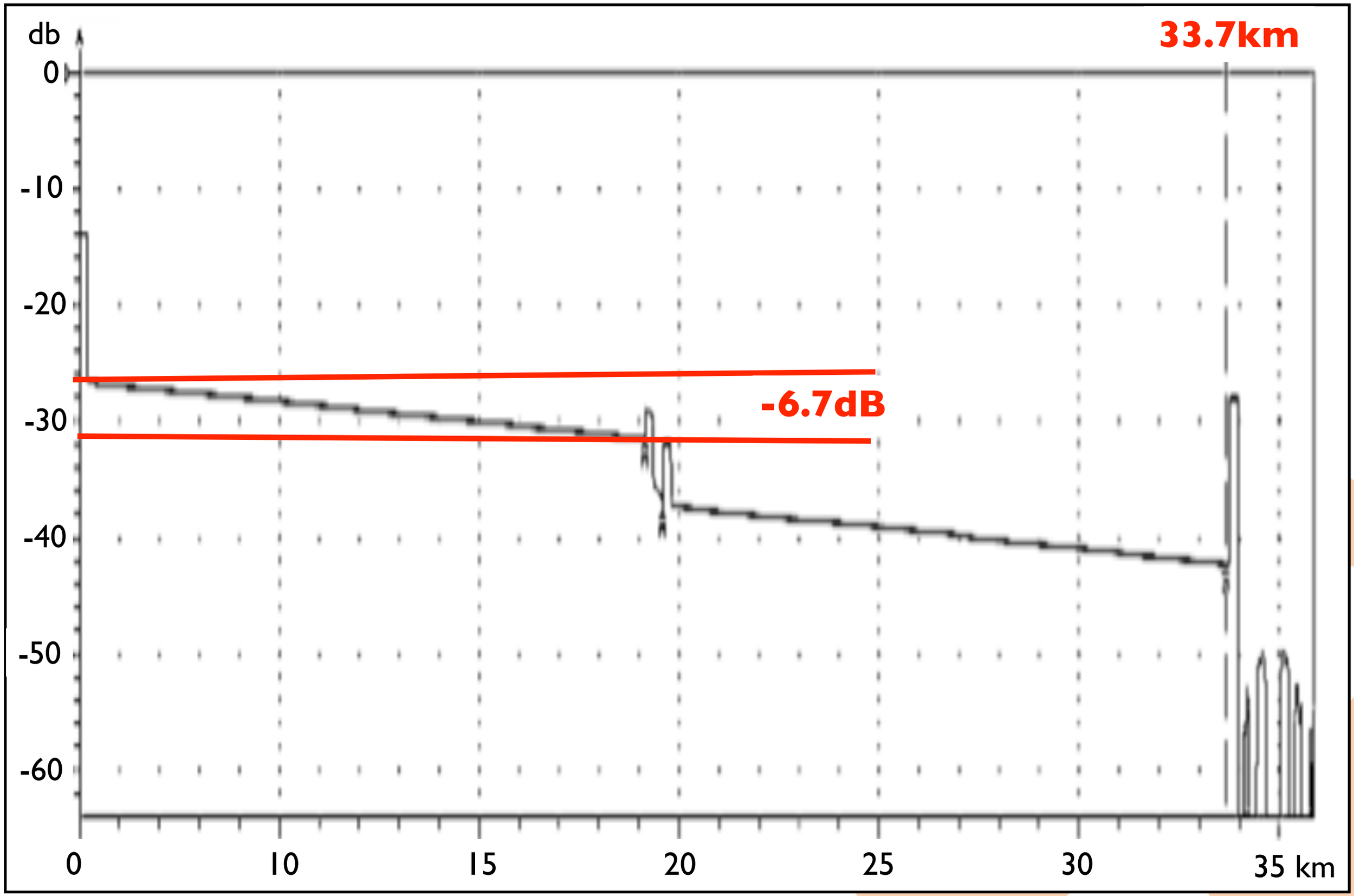
Problem Statement

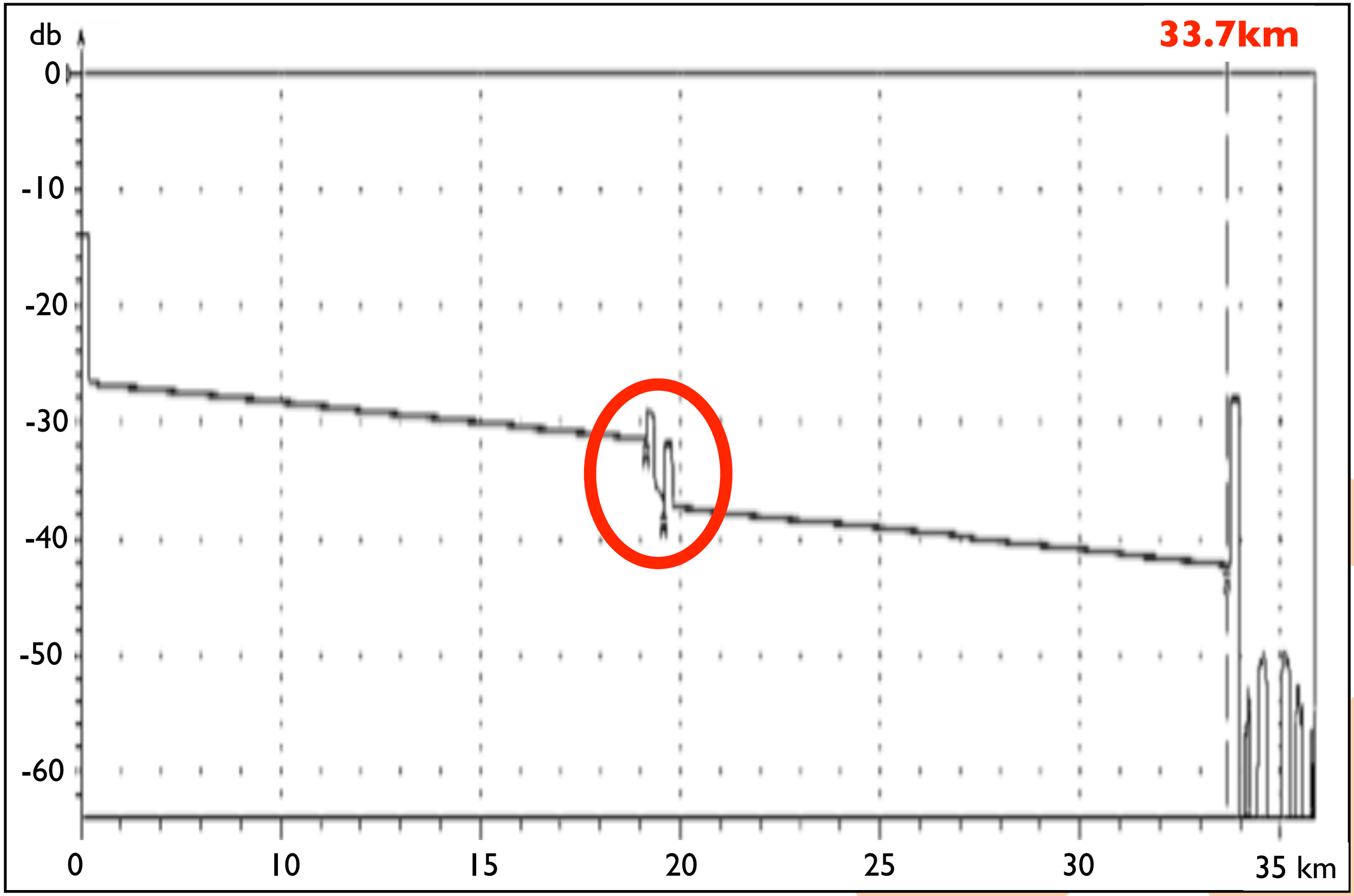
- We need to extend the range of currently available 100Gbit/s technology.
- What are the options?
  - Amplification
  - Transmission equipment
- What do we have to deal with
  - Attenuation
  - Dispersion

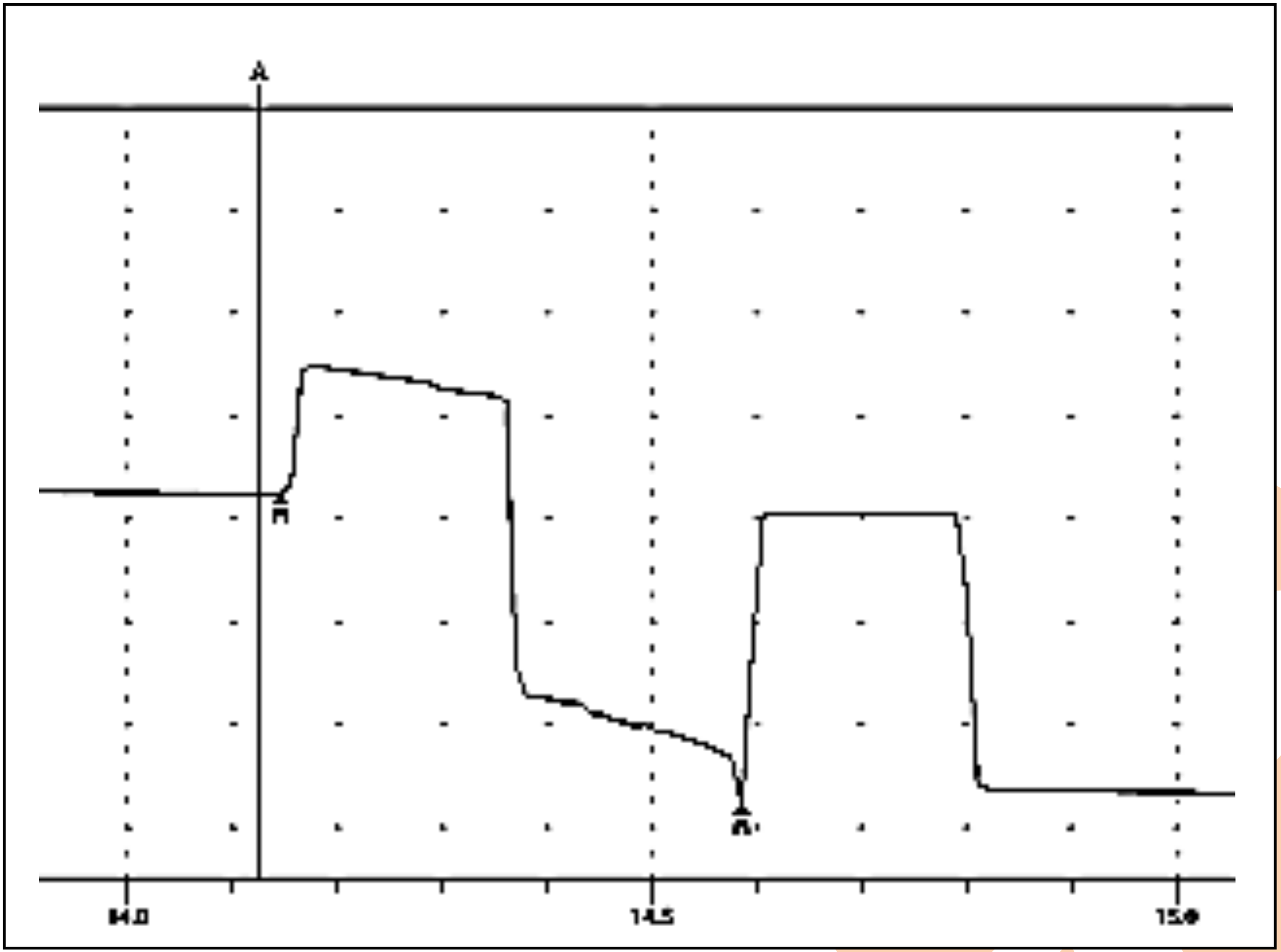

3
2.5
2
1.5
1
0.5
0
Attenuarion (dB/km)
0.35dB/km
0.22dB/km
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavelength (nm)

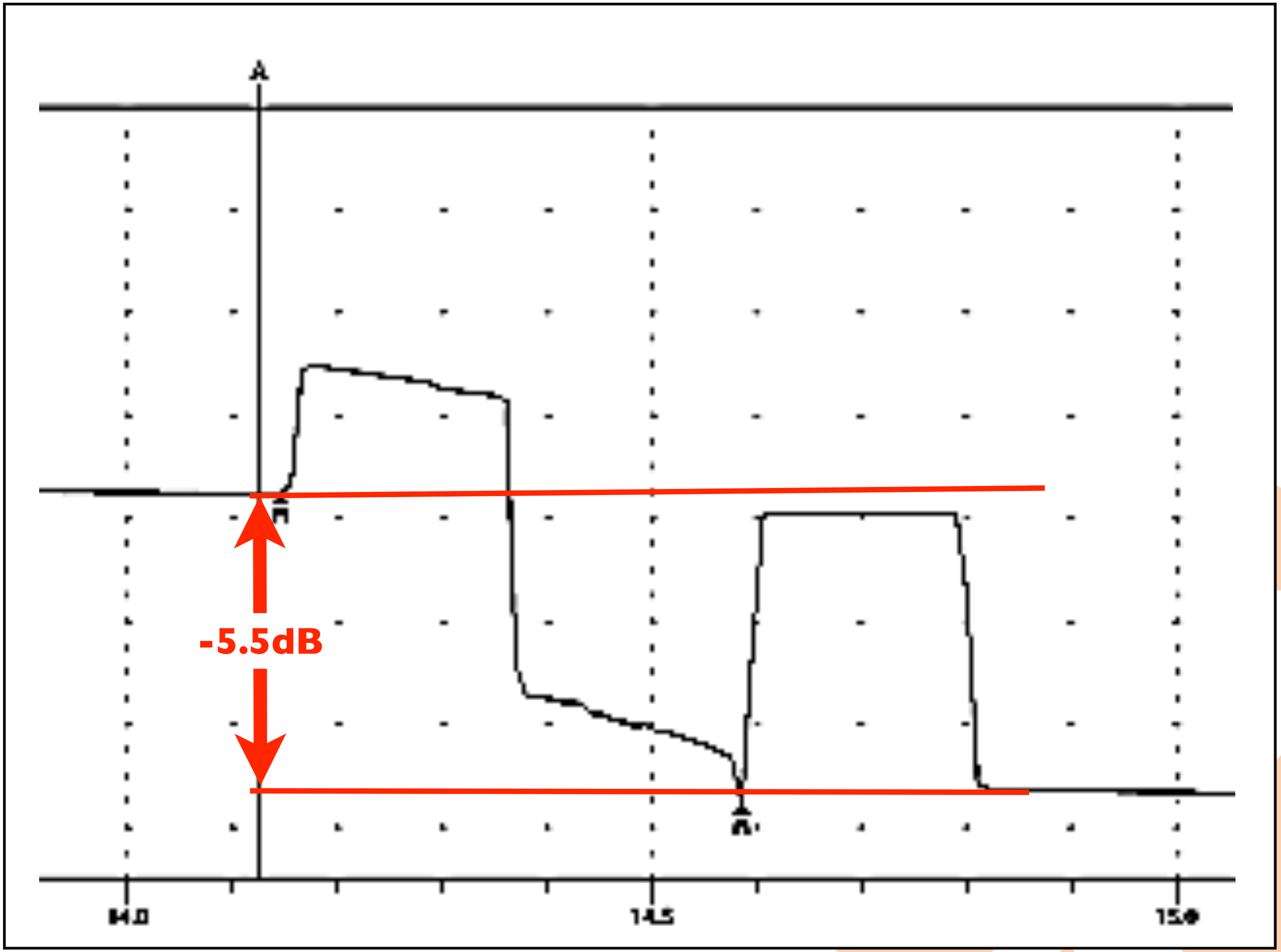

db
0
-10
-20
-30
-40
-50
-60
0
10
15
20
25
30
35 km


db
0
-10
-20
-30
-40
-50
-60
33.7km
-20dB
0
10
15
20
25
30
35 km

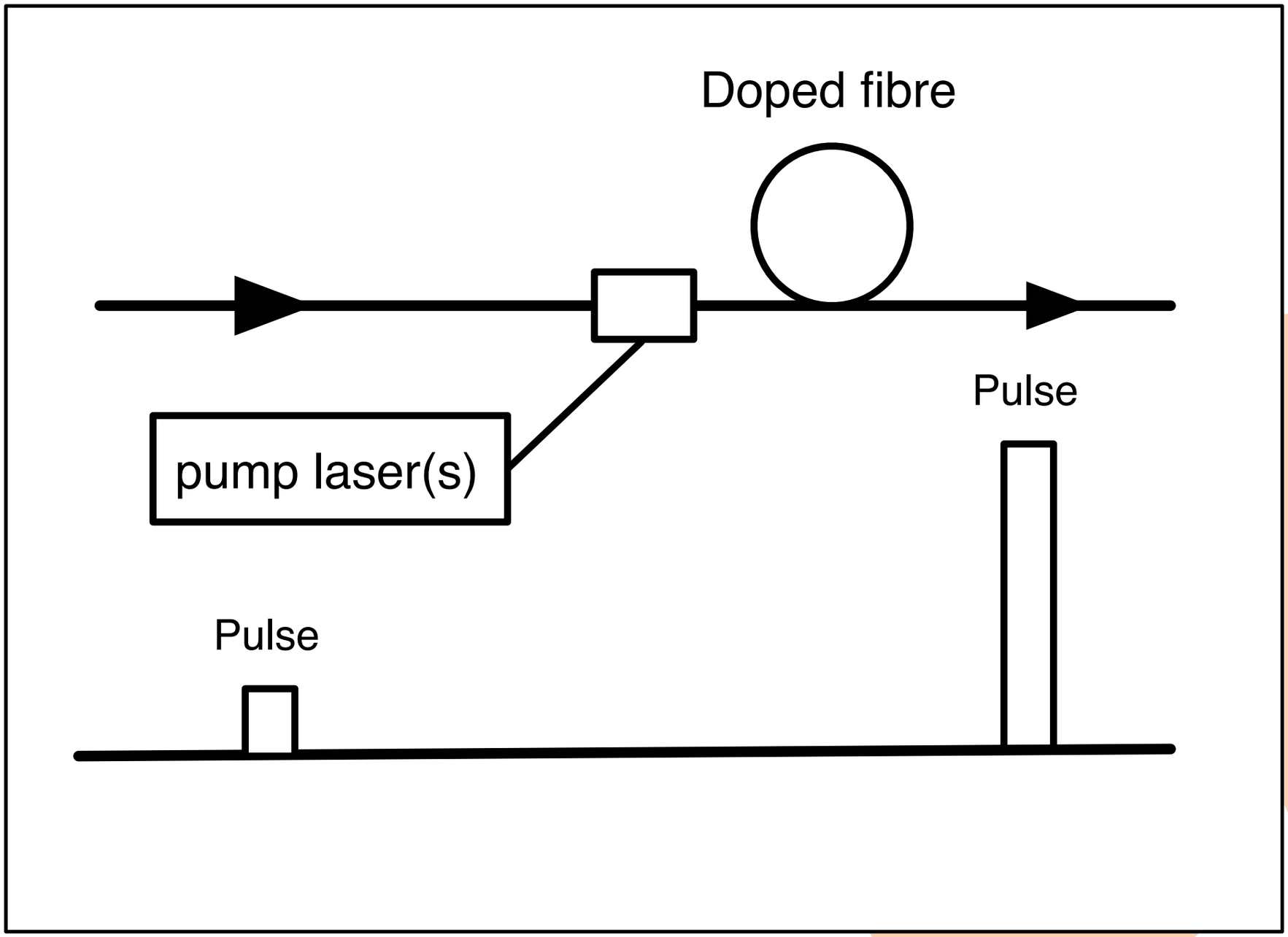

db
0
-10
-20
-30
-40
-50
-60
33.7km
-6.7dB
0
10
15
20
25
30
35 km


db
0
-10
-20
-30
-40
-50
-60
33.7km
0
10
15
20
25
30
35 km

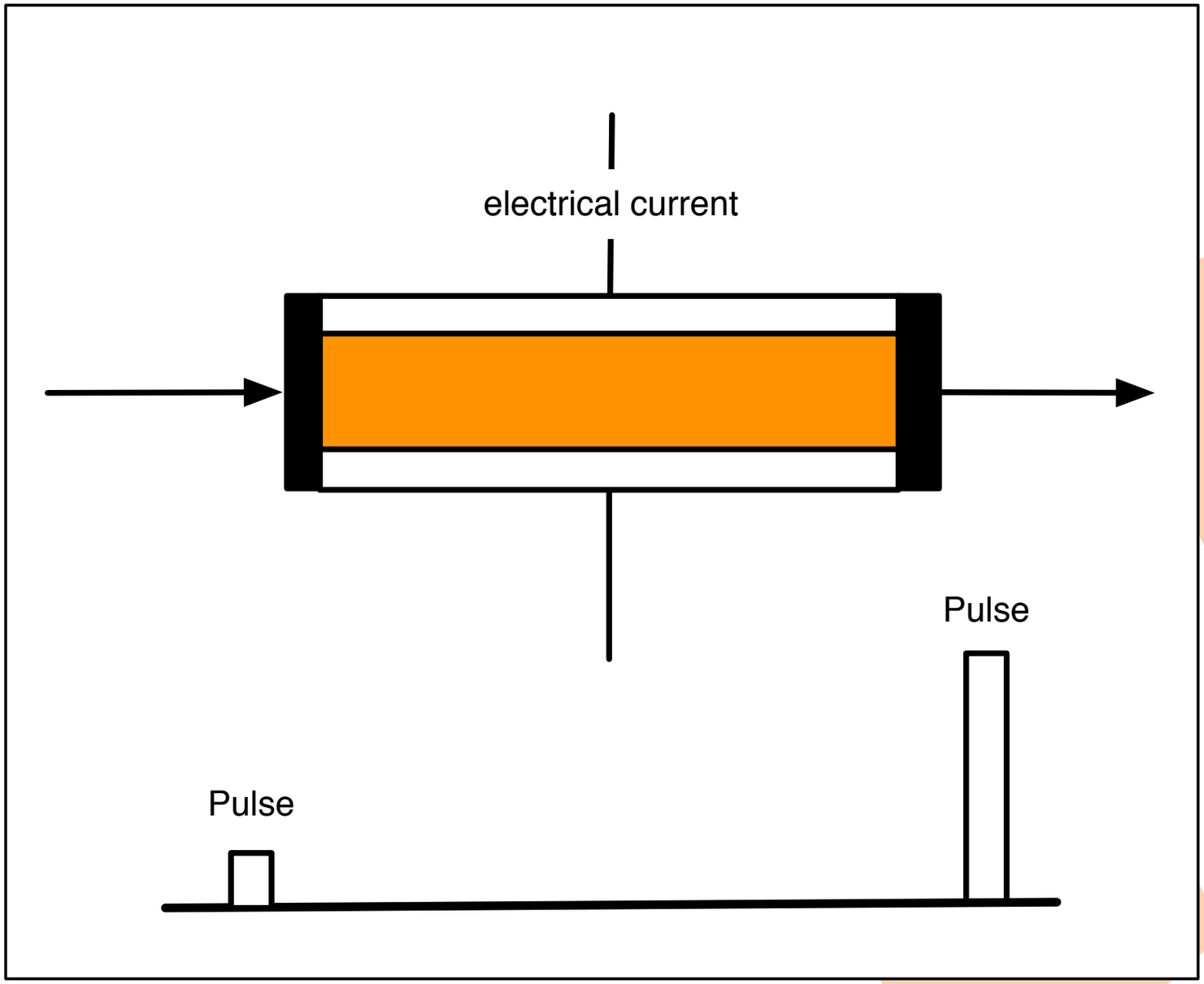

A
B
C
14.0
14.5
15.0


A
-5.5dB
14.0
14.5
15.0

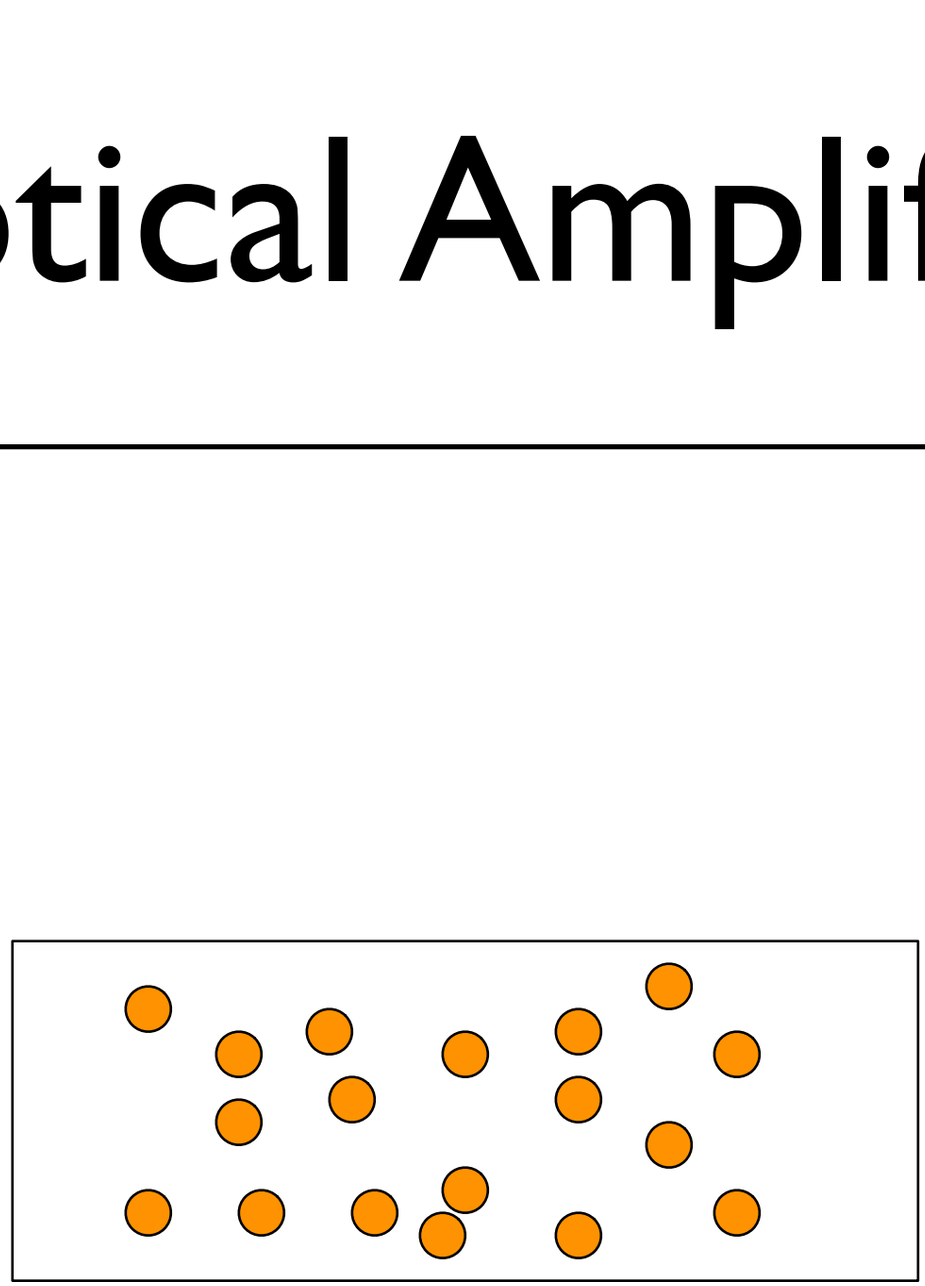
Doped Fibre Amplifier



Semiconductor Optical Amplifier



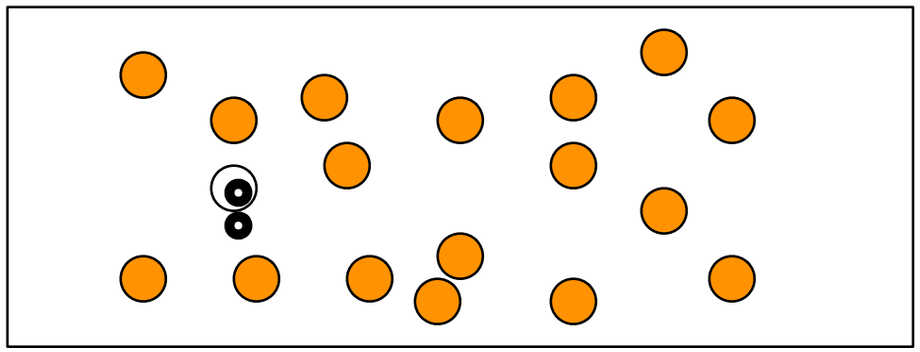
Optical Amplifier



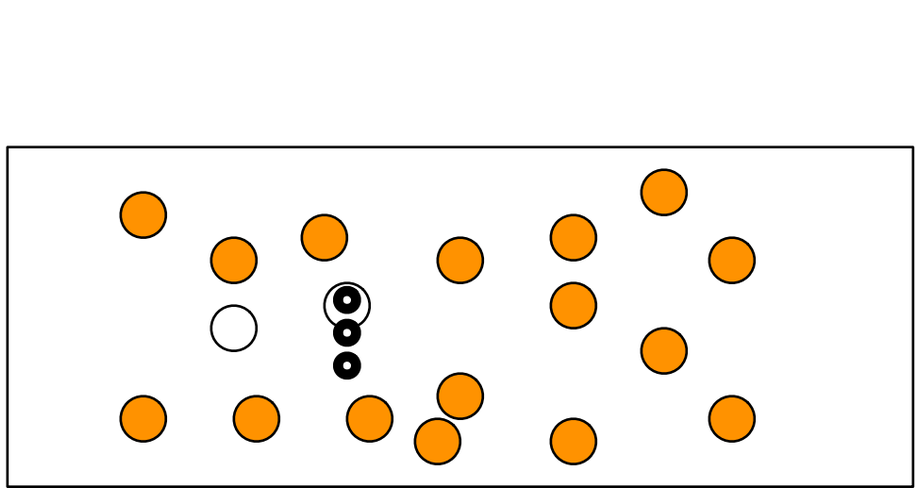
Optical Amplifier

Optical Amplifier

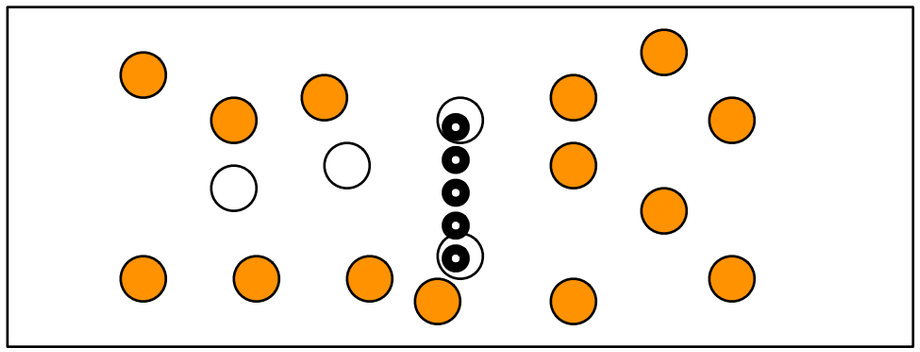
Optical Amplifier



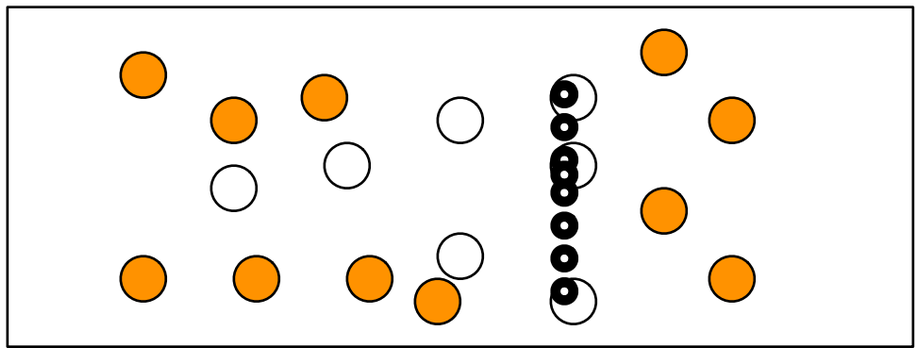
Optical Amplifier



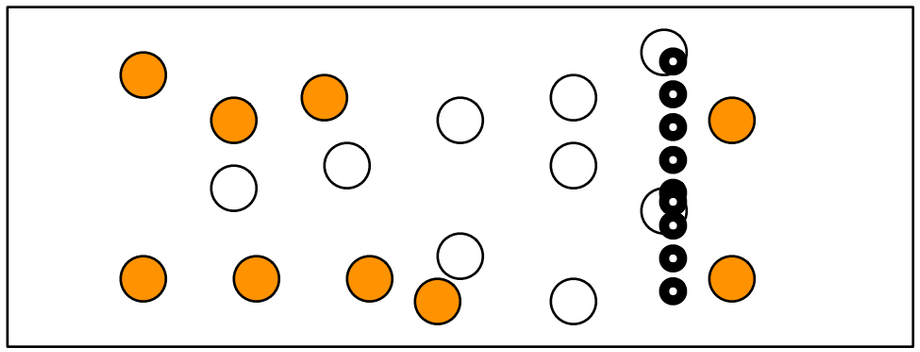
Optical Amplifier



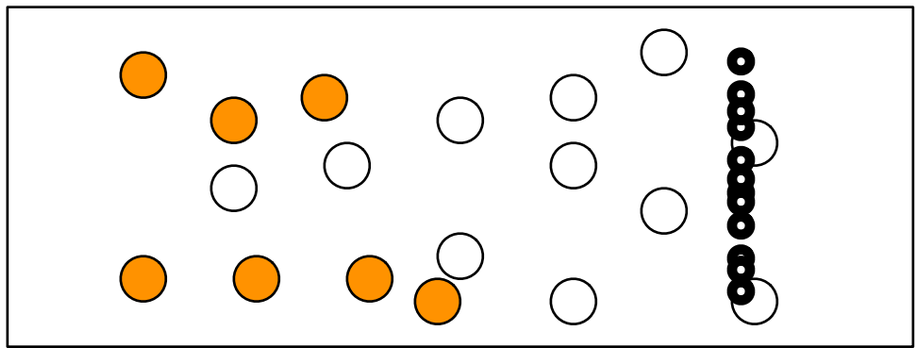
Optical Amplifier



Optical Amplifier

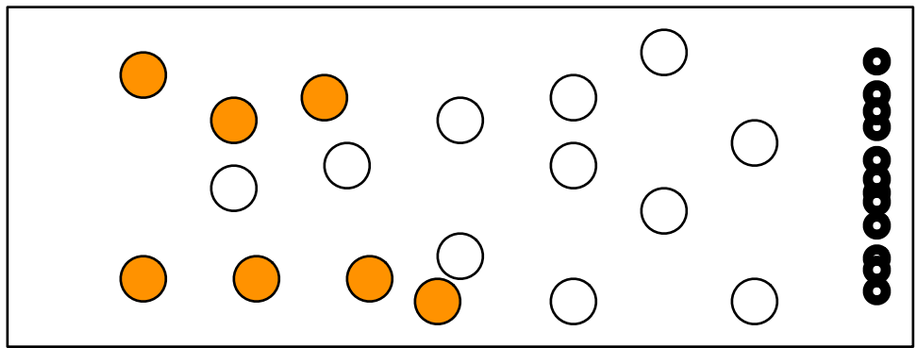


Optical Amplifier

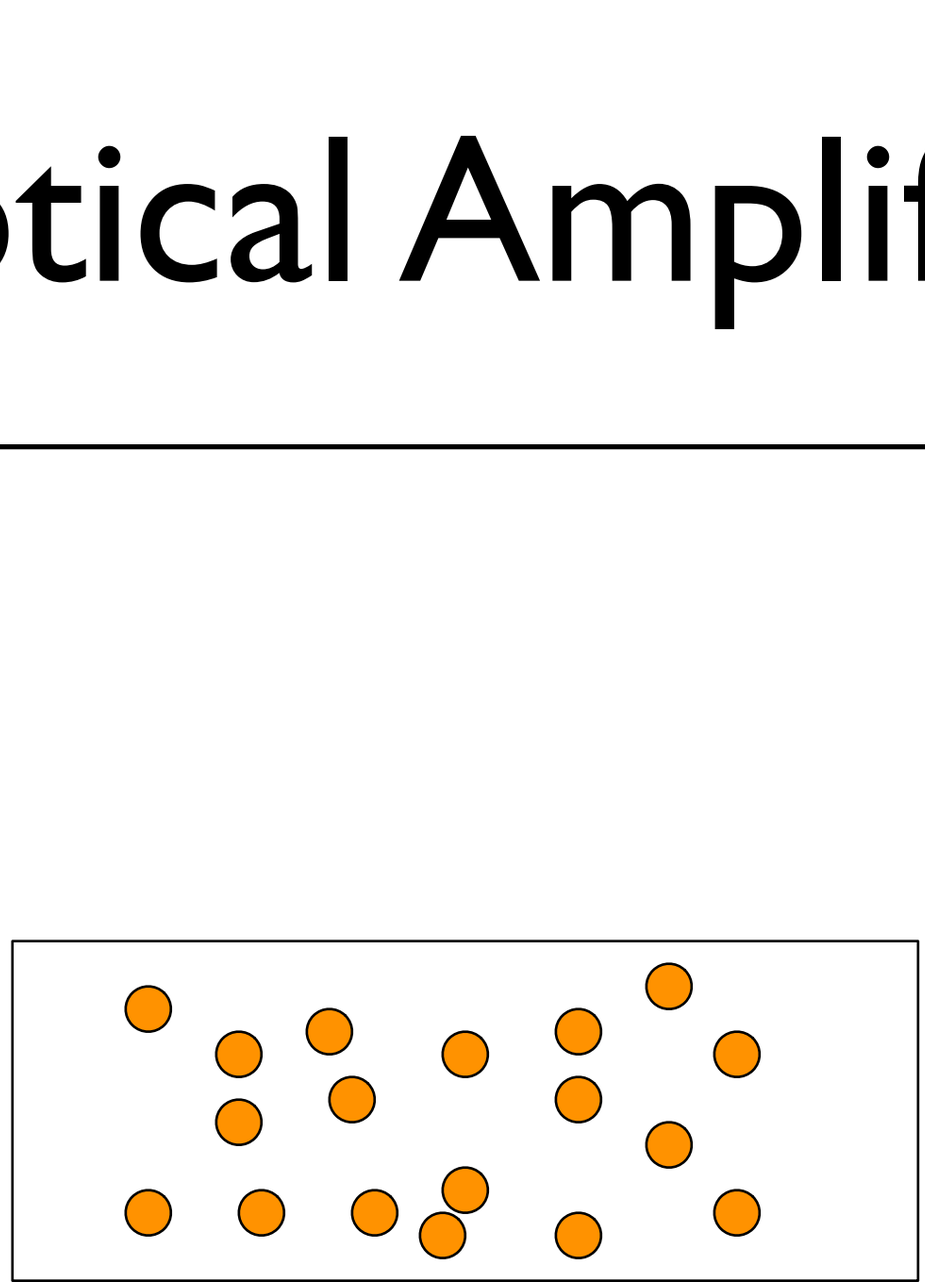


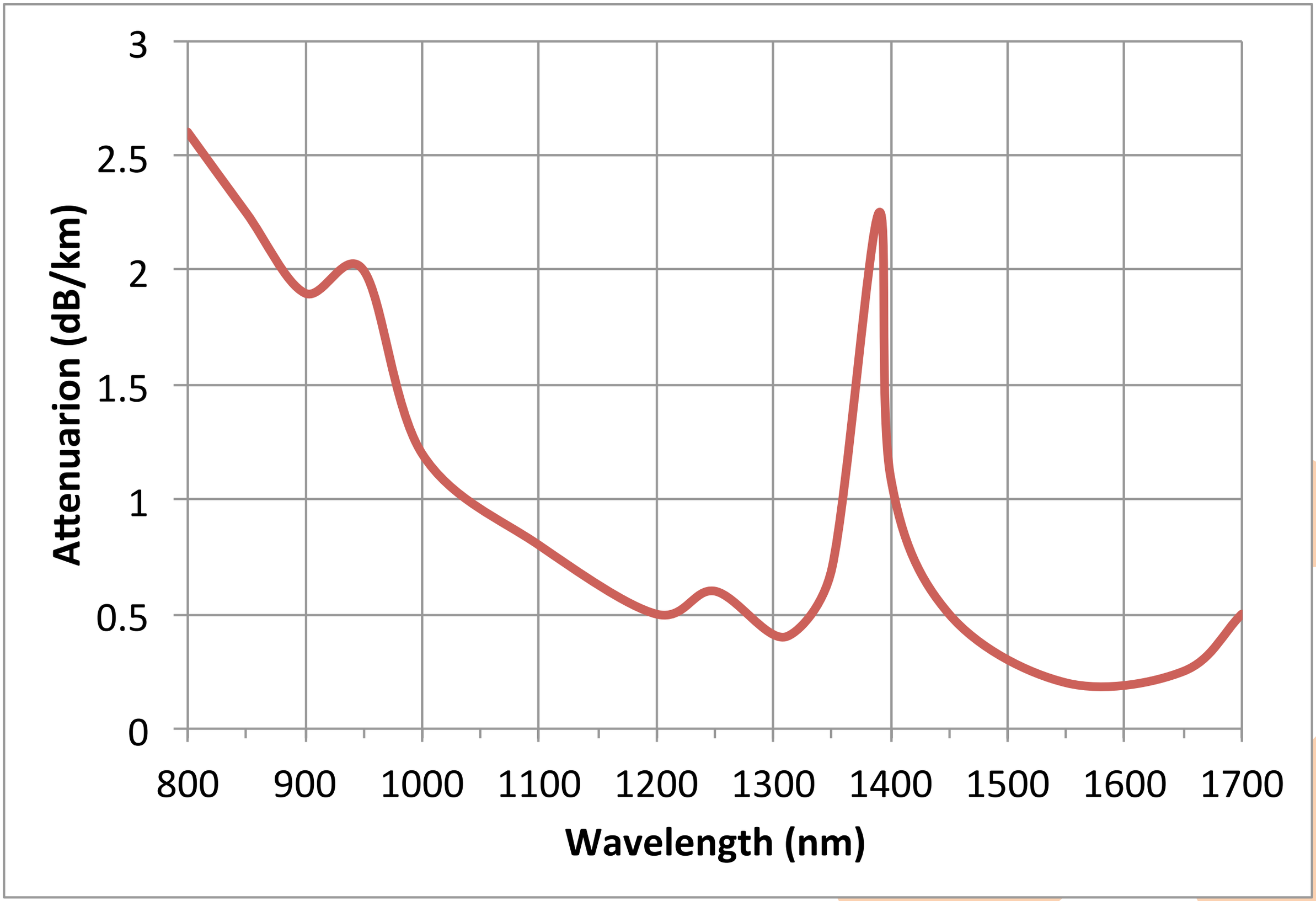
Optical Amplifier

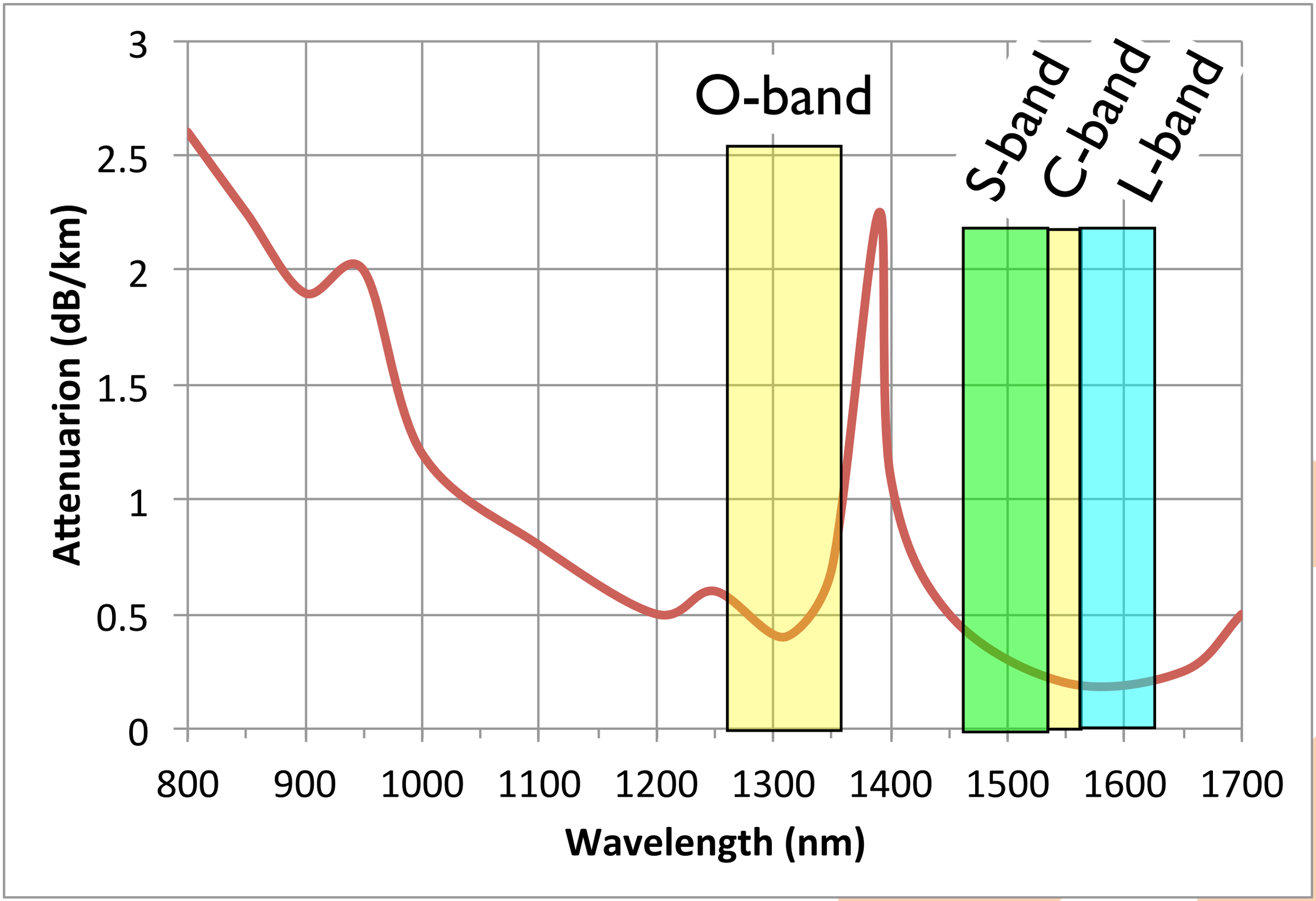
Optical Amplifier



Optical Amplifier

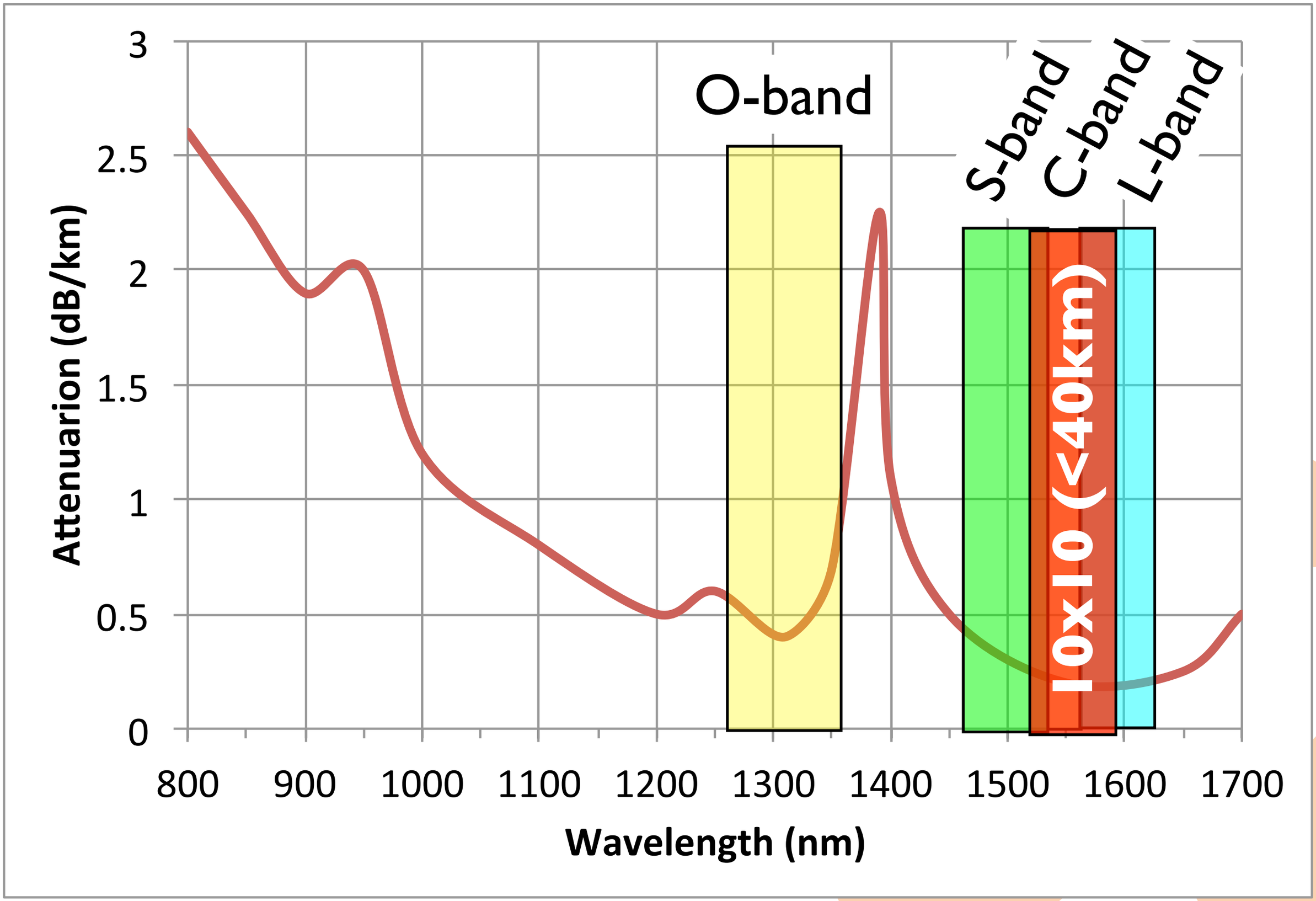



3
2.5
2
1.5
1
0.5
0
Attenuarion (dB/km)
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavelength (nm)


O-band
S-band
C-band
L-band
Attenuation (dB/km)
3
2.5
2
1.5
1
0.5
0
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavelength (nm)

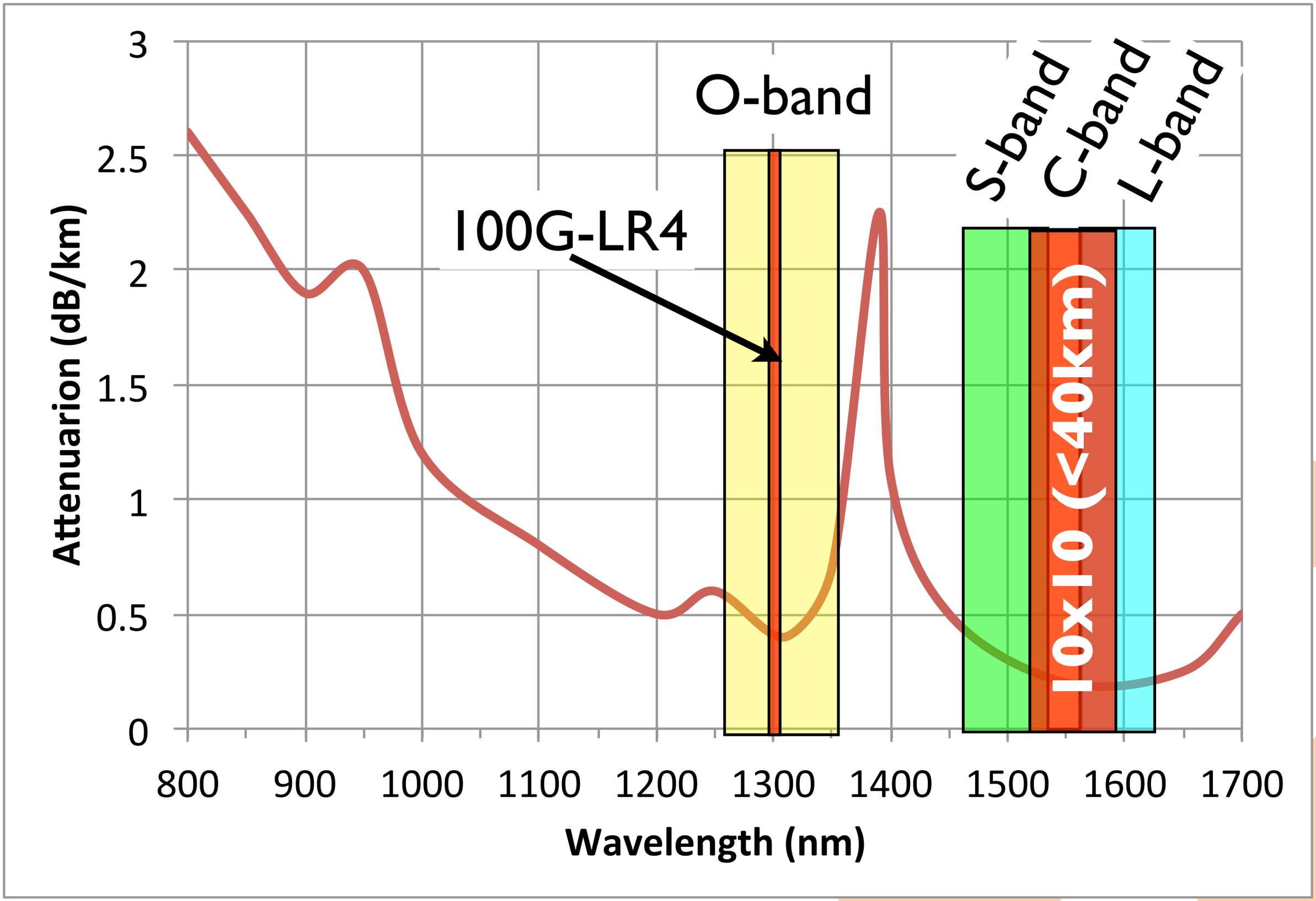
Band	Doped Fibre Amplifier
Original (O) 1260 to 1360 nm	PDFA (Praseodymium doped fibre amplifier)
Short wavelengths (S) 1460 to 1530 nm	TDFA (Thulium doped fibre amplifier)
Conventional (C) 1530 to 1565 nm	EDFA (C-band) (Erbium doped fibre amplifier)
Long wavelengths (L) 1565 to 1625 nm	EDFA (L-band) (Erbium doped fibre amplifier)

Lane	Center λ (nm)	λ Range (nm)
L_0	1523	1521 to 1525
L_1	1531	1529 to 1533
L_2	1539	1537 to 1541
L_3	1547	1545 to 1549
L_4	1555	1553 to 1557
L_5	1563	1561 to 1565
L_6	1571	1569 to 1573
L_7	1579	1577 to 1581
L_8	1587	1585 to 1589
L_9	1595	1593 to 1597

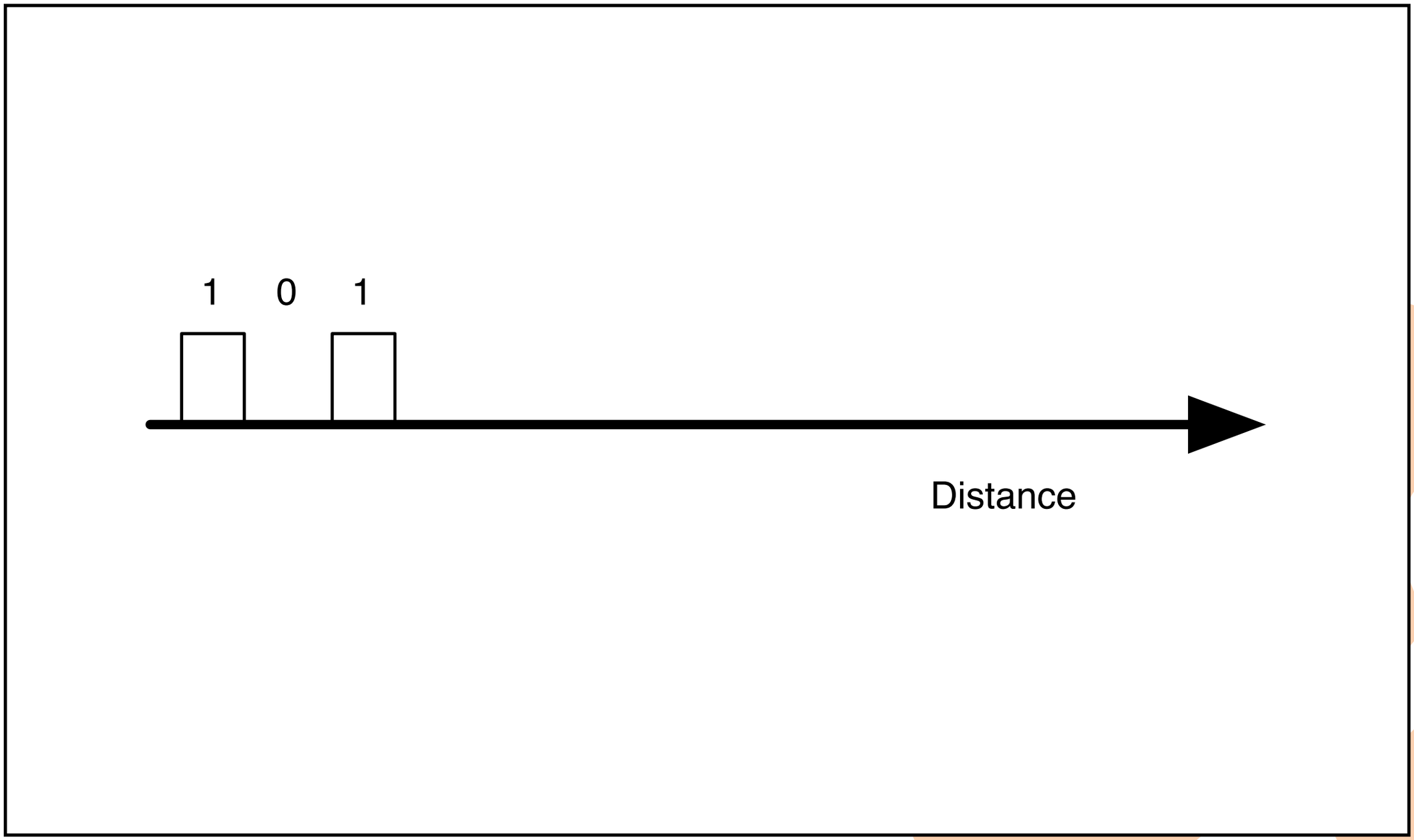

O-band
S-band
C-band
L-band
10x10 (<40km)
Attenuation (dB/km)
3
2.5
2
1.5
1
0.5
0
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavelength (nm)

100G-LR4

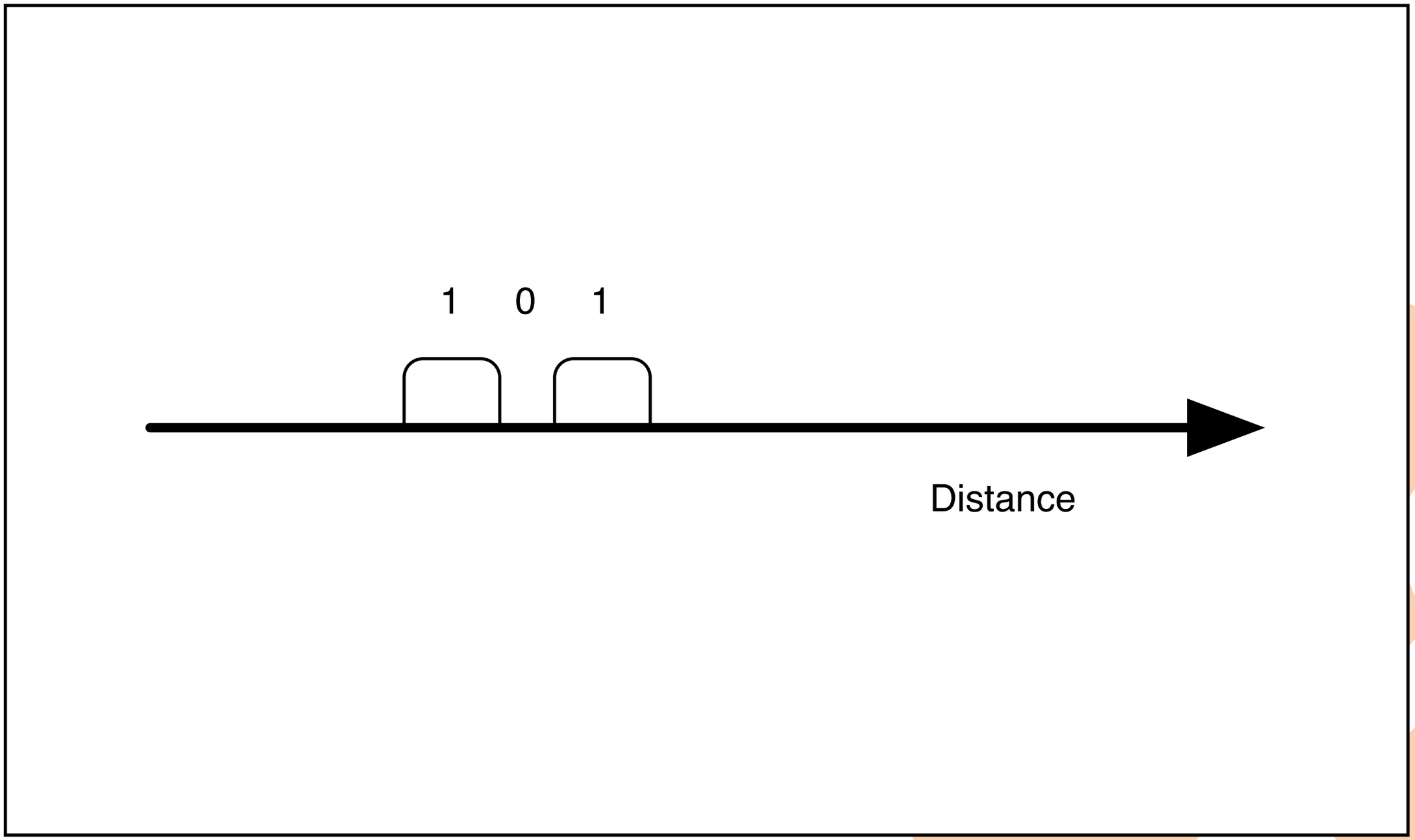
Lane	Center λ (nm)	λ Range (nm)
L_0	1295.56	1294.53 to 1296.59
L_1	1300.055	1299.02 to 1301.09
L_2	1304.585	1303.54 to 1305.63
L_3	1309.14	1308.09 to 1310.19


O-band
100G-LR4
S-band
C-band
L-band
10x10 (<40km)
Attenuation (dB/km)
3
2.5
2
1.5
1
0.5
0
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavelength (nm)

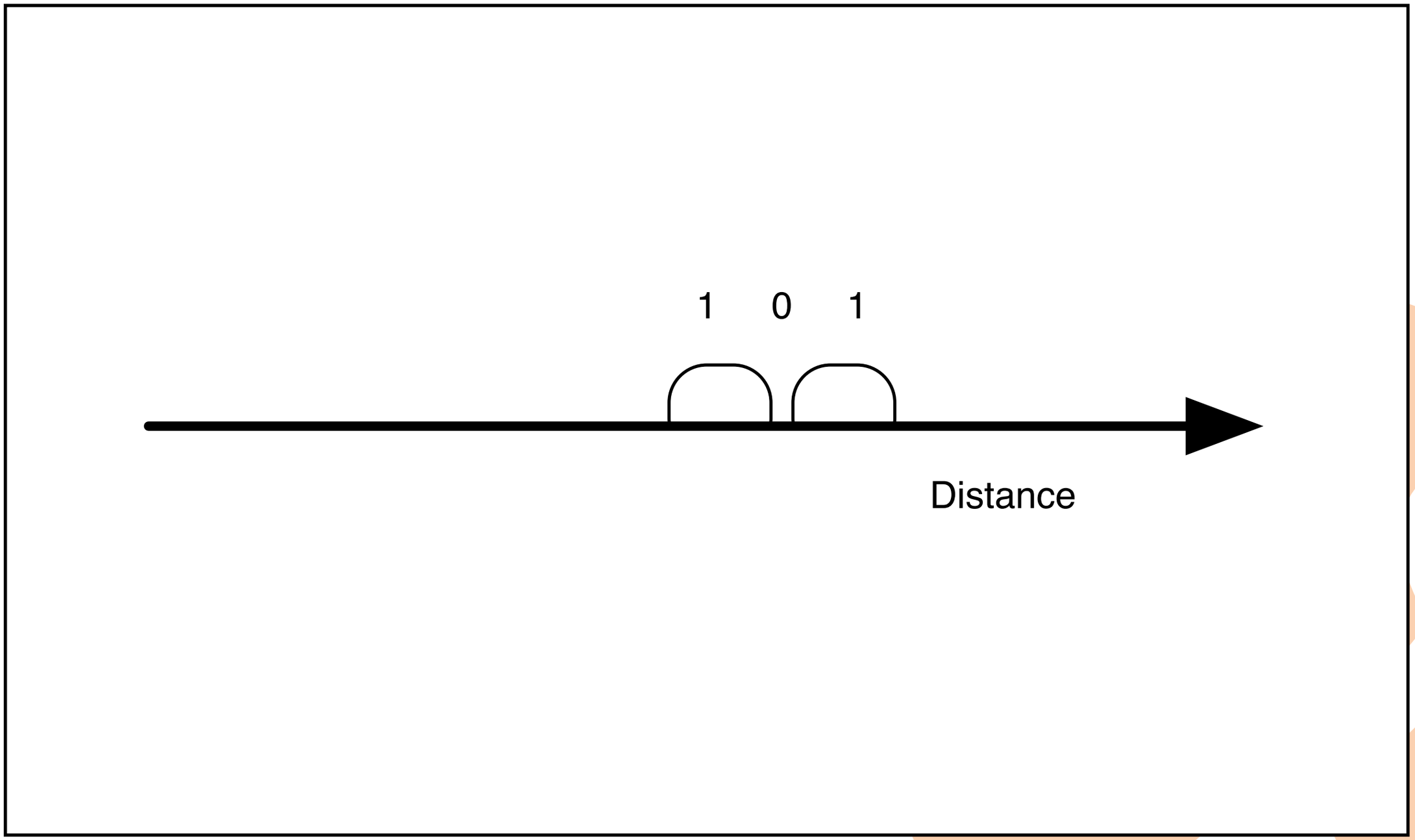
Dispersion



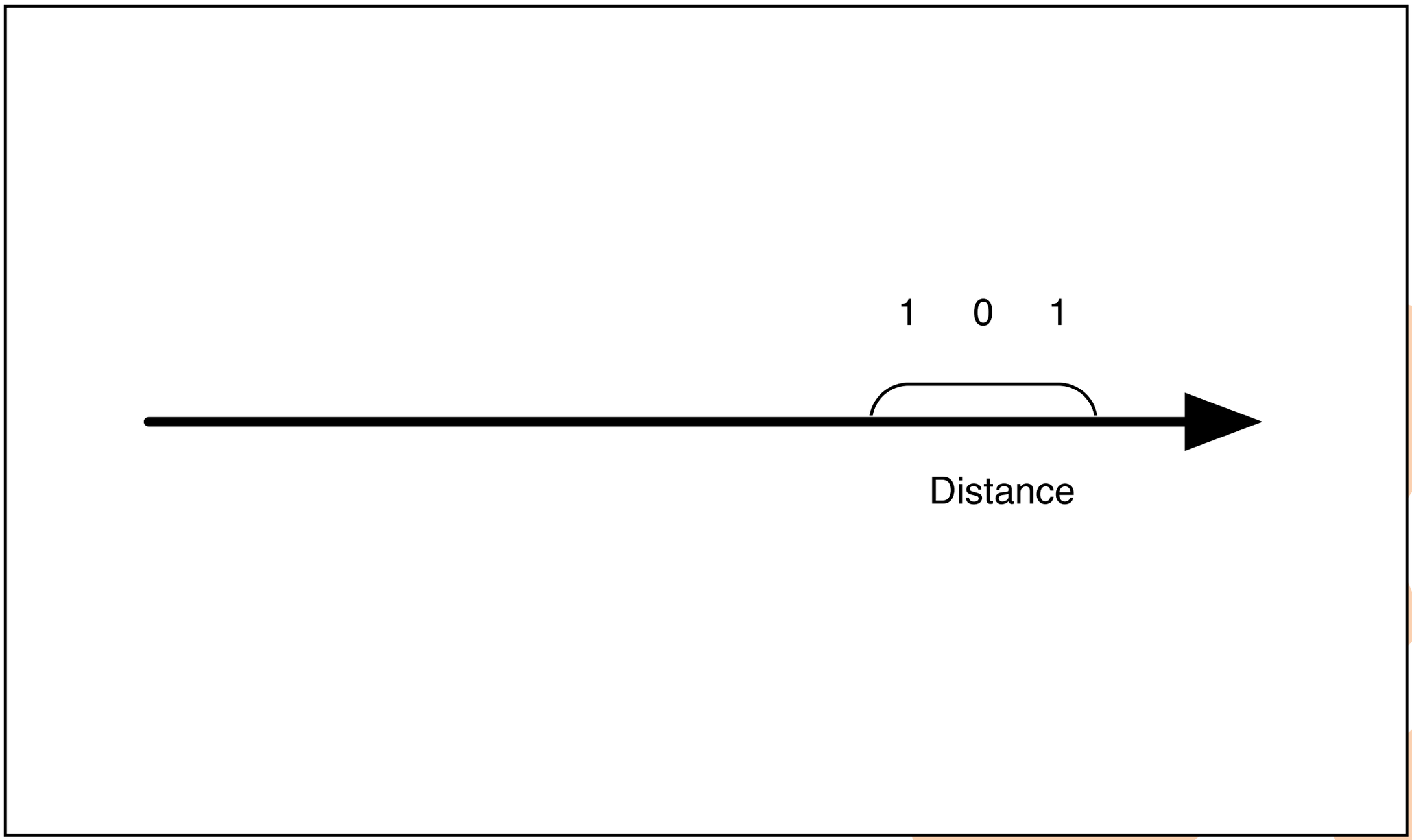
Dispersion



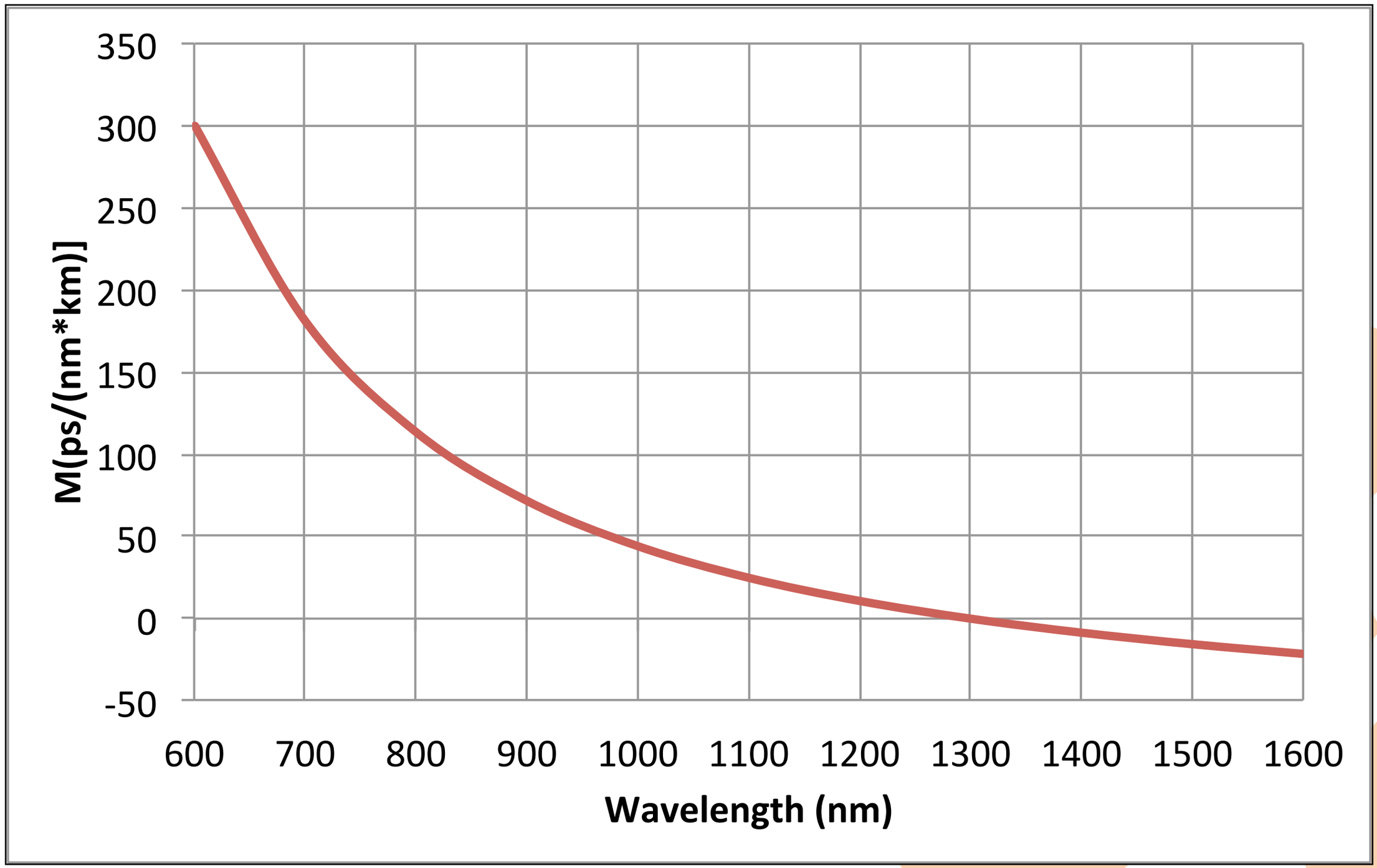
Dispersion



Dispersion

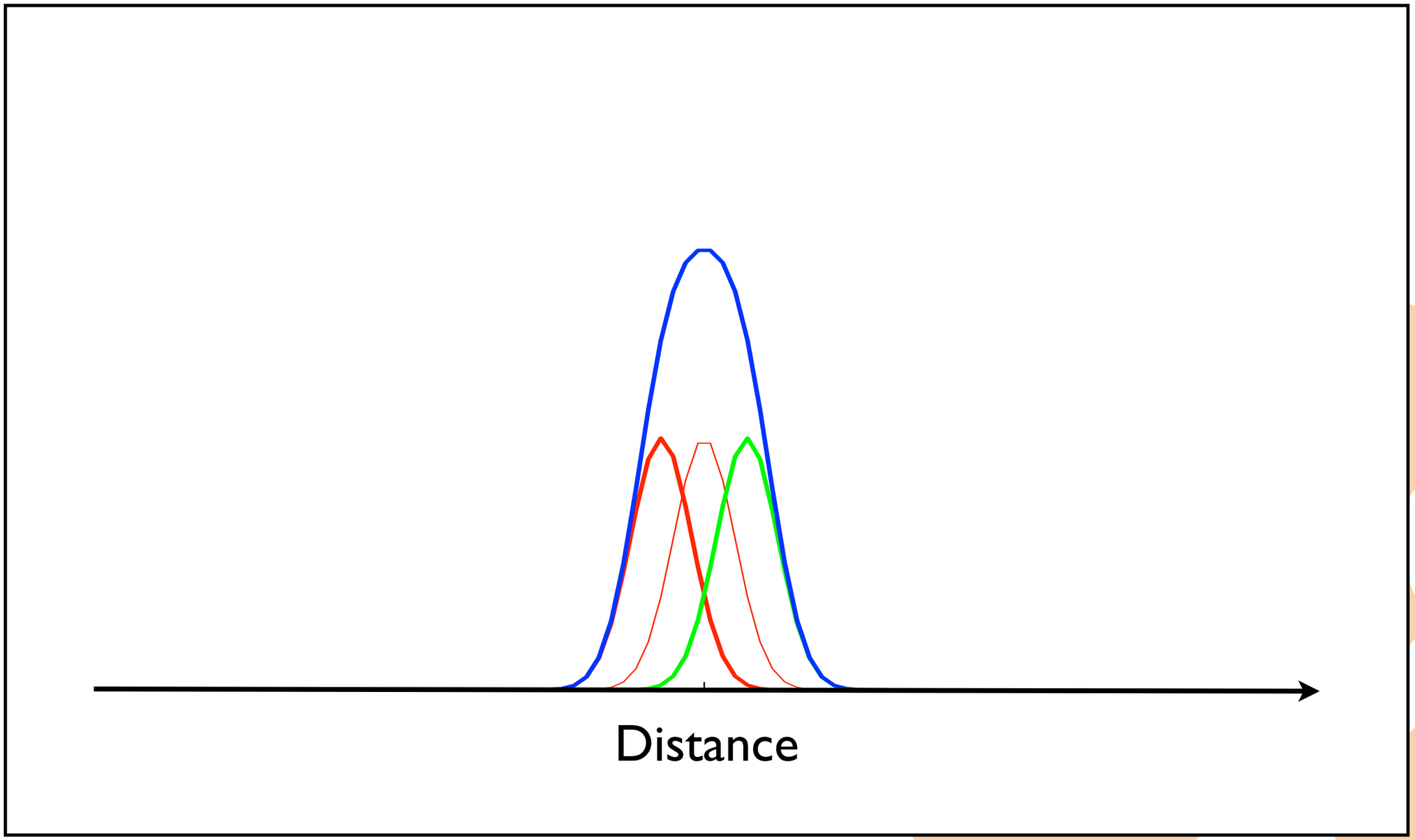


Material Dispersion

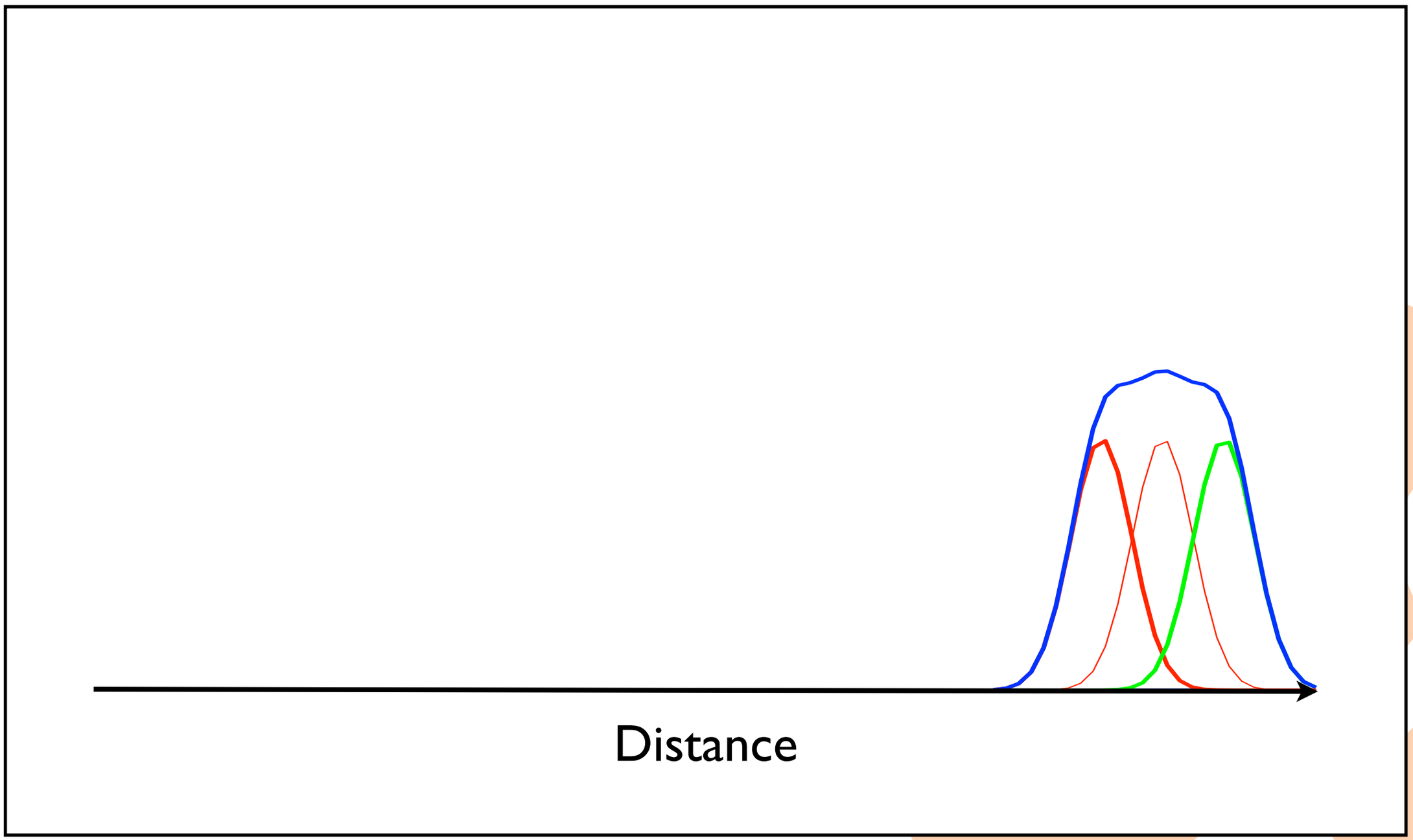


Material Dispersion

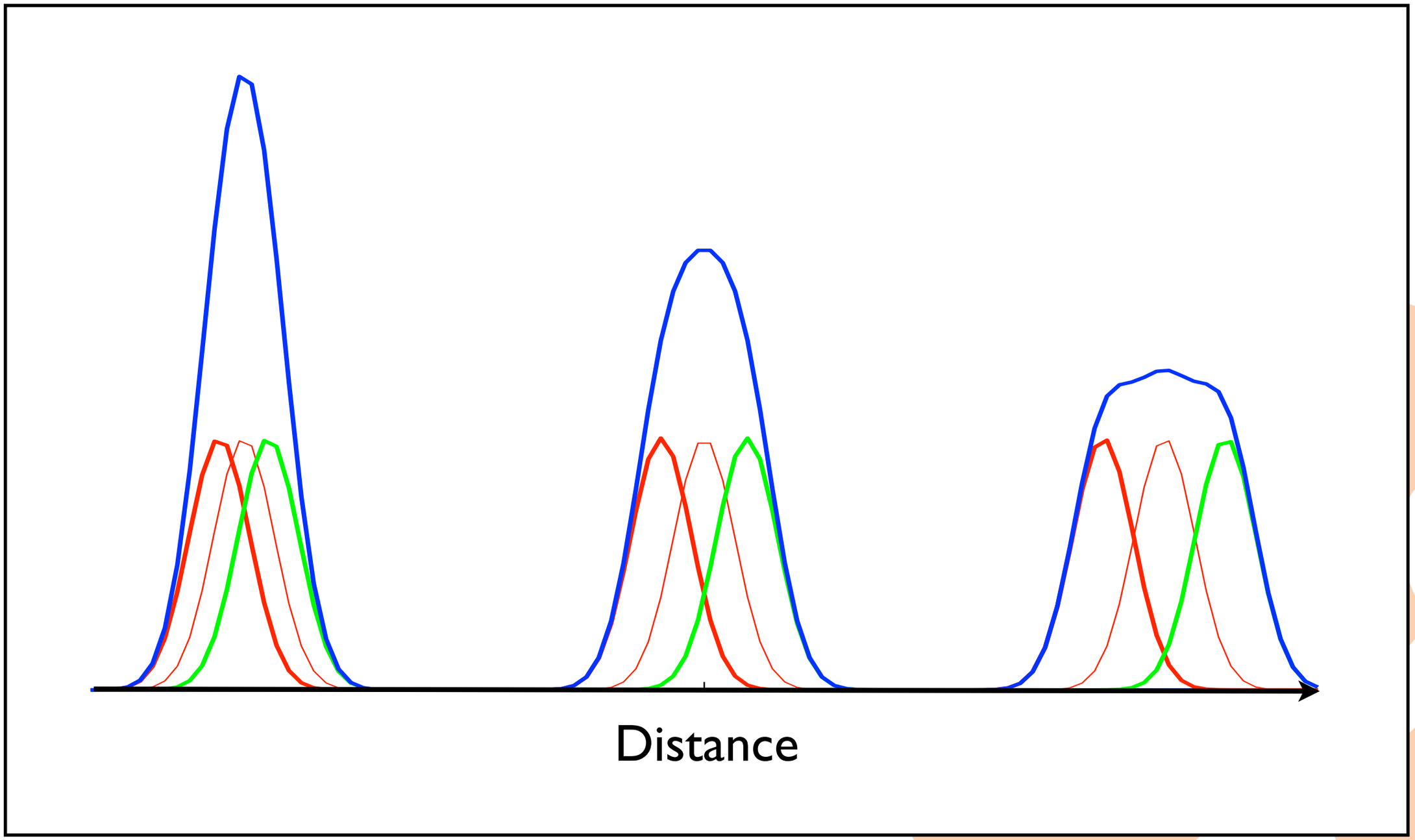
Material Dispersion



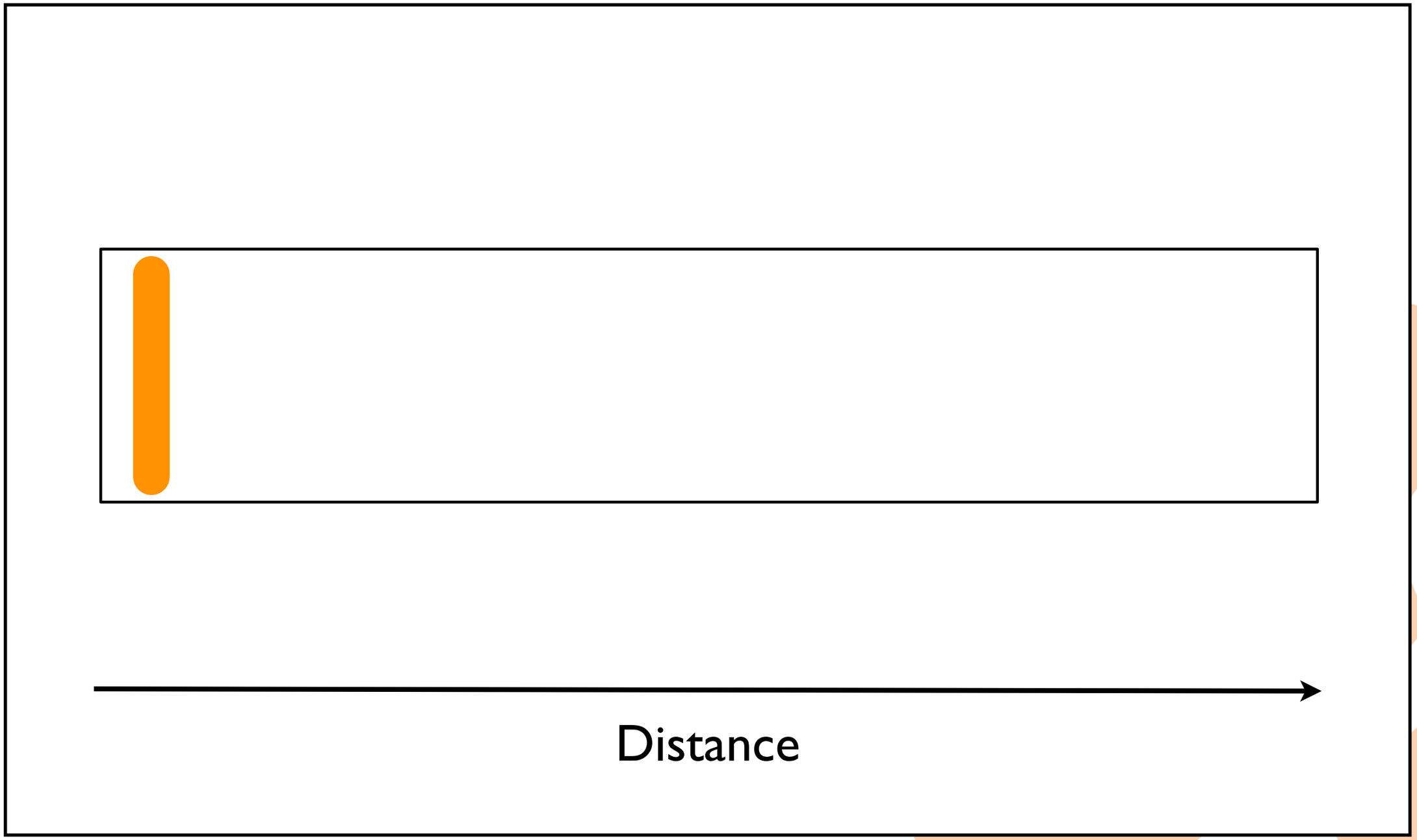
Material Dispersion



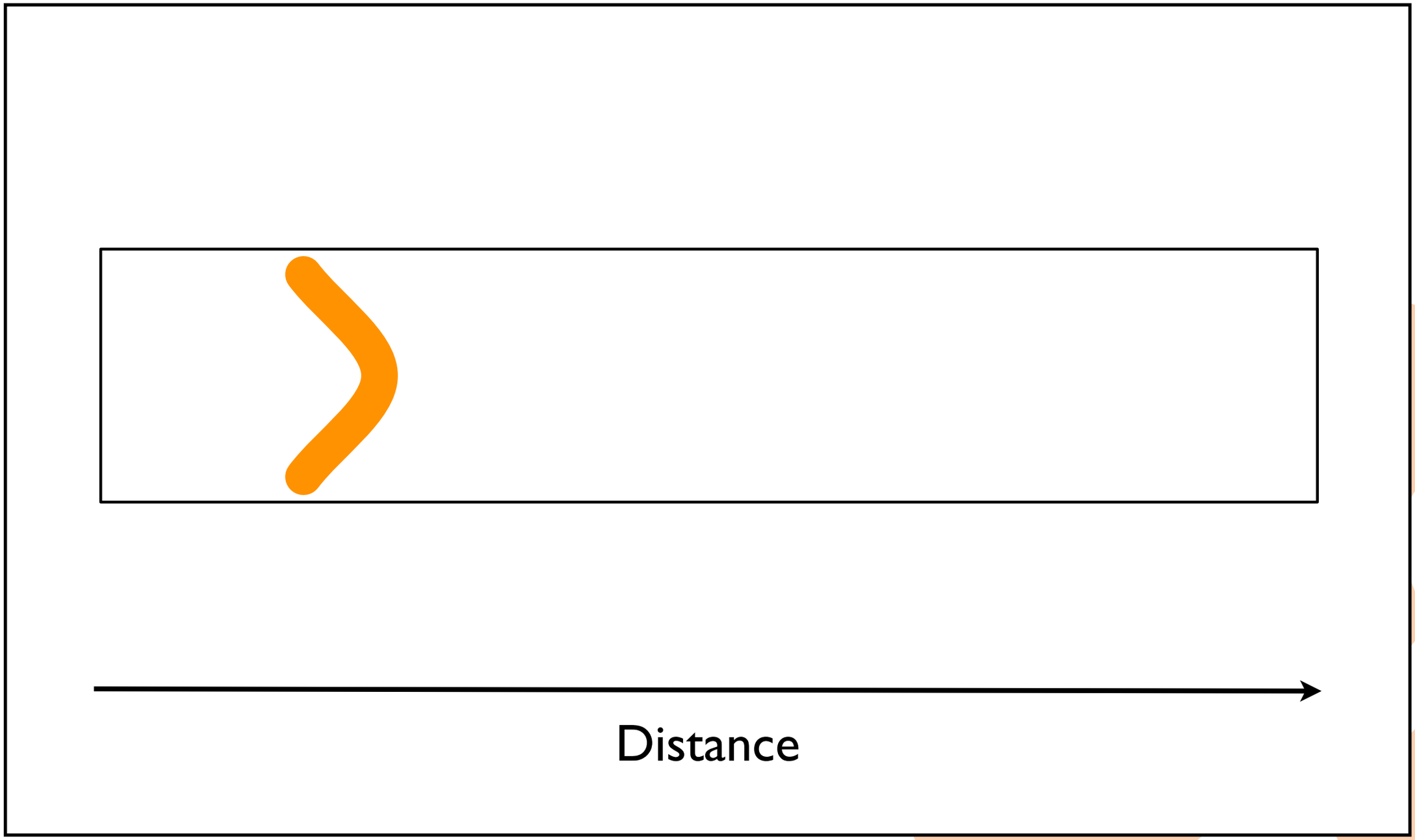
Material Dispersion



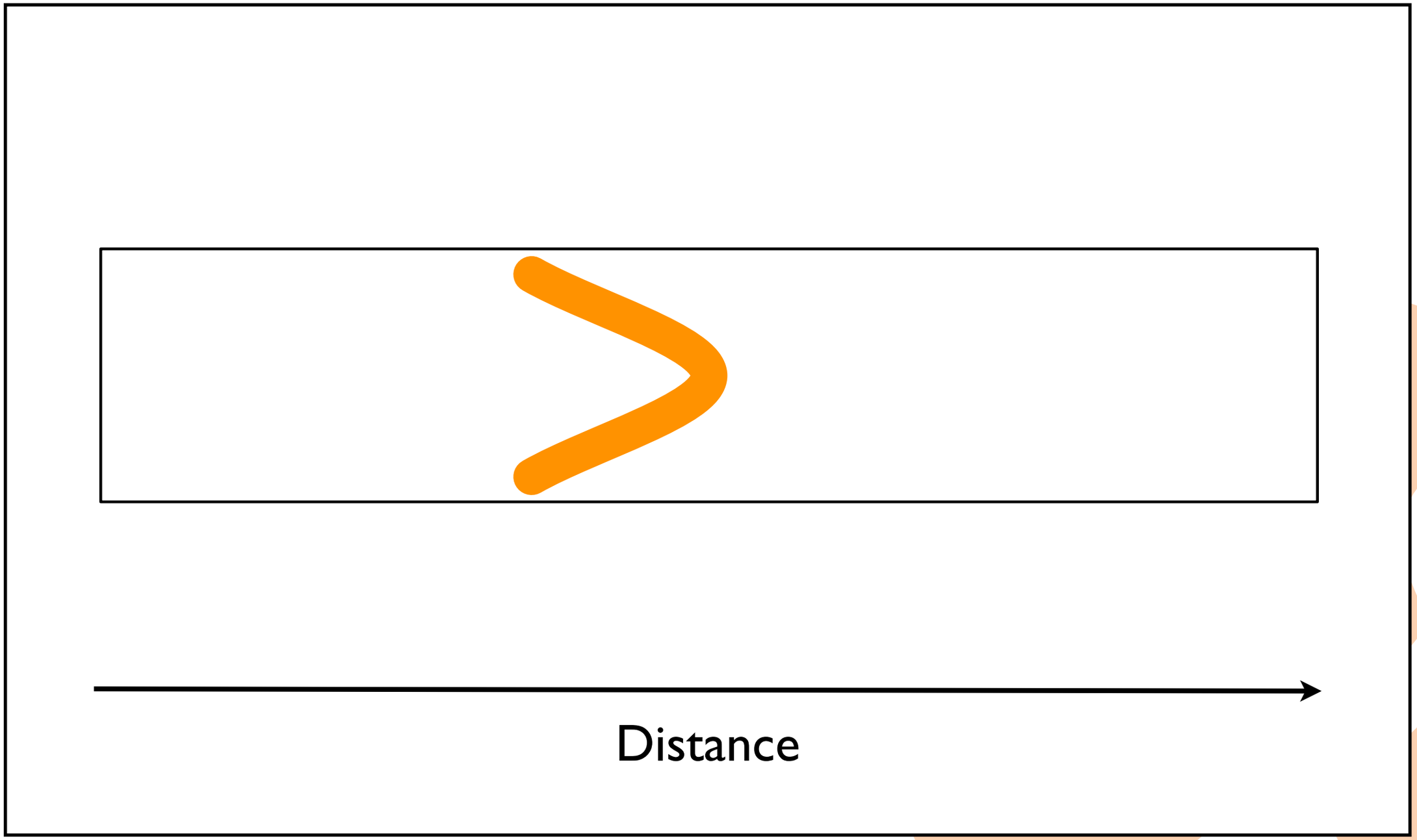
Waveguide Dispersion



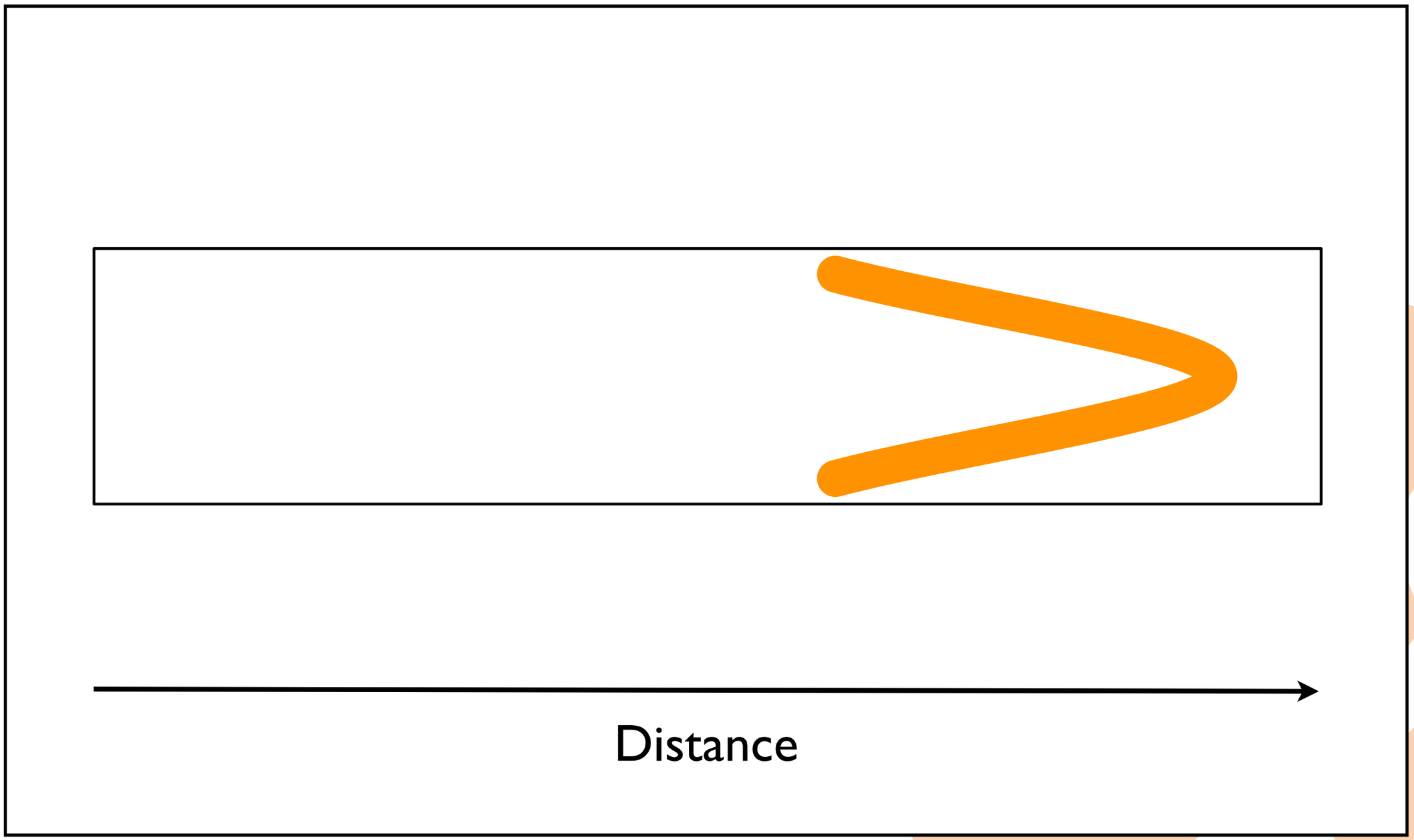
Waveguide Dispersion



Waveguide Dispersion



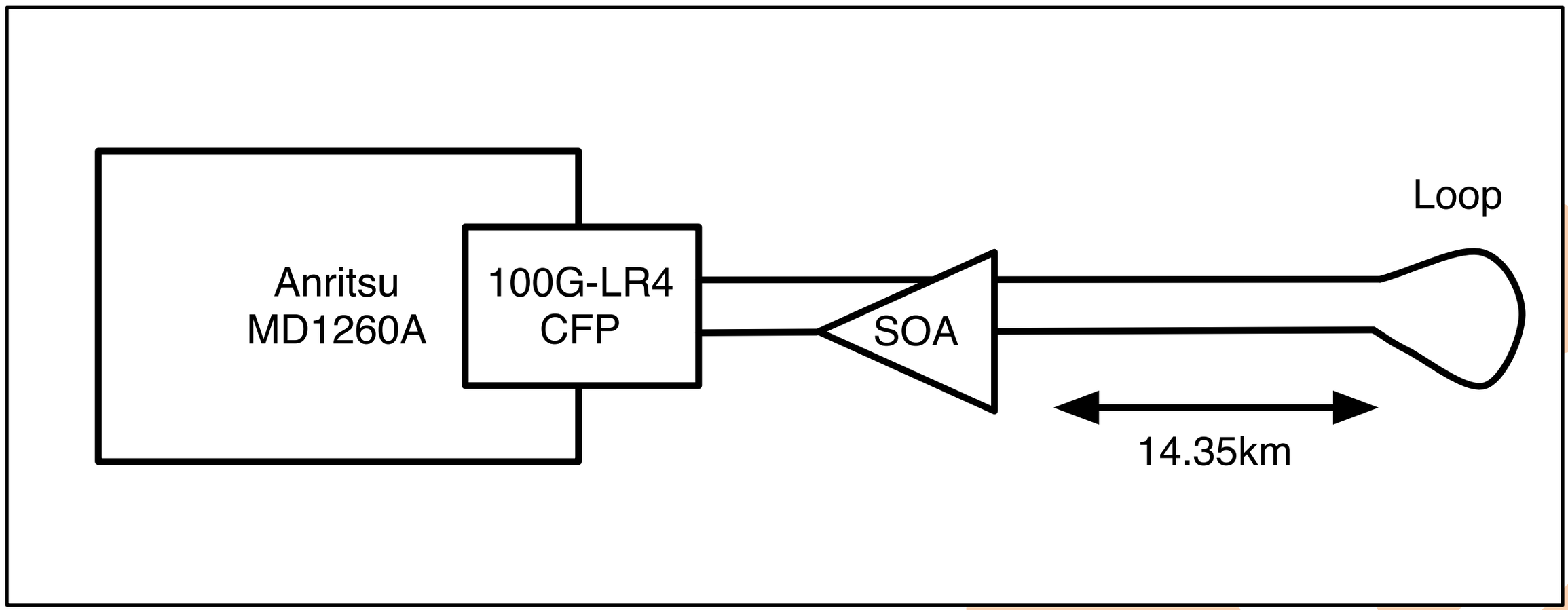
Waveguide Dispersion



Options

- LR-4 (O-band) is the only option.
  - Only type (band) that can be amplified.
  - 0 material dispersion.
- Semiconductor Optical Amplifiers (SOAs)
  - As component available for all bands.
  - Much much cheaper than PDFAs.
  - We had one for the O-Band from 2005.

Experiment



Result

- 6 trillion, 500 billion frames (6,500,000,000,000)
- Receives, all - 2 (6,499,999,999,998)

Conclusion Nov. 2011

- Doped Fibre Amplifiers only suitable for O-band.
  - PDFA
  - Very expensive but we'll bought two to gain experience.
- Semiconductor Optical Amplifier.
  - Seems to work well for the O-band.
  - No ready made appliance
    (it's a component, just like in 2005).

Next steps Nov. 2011

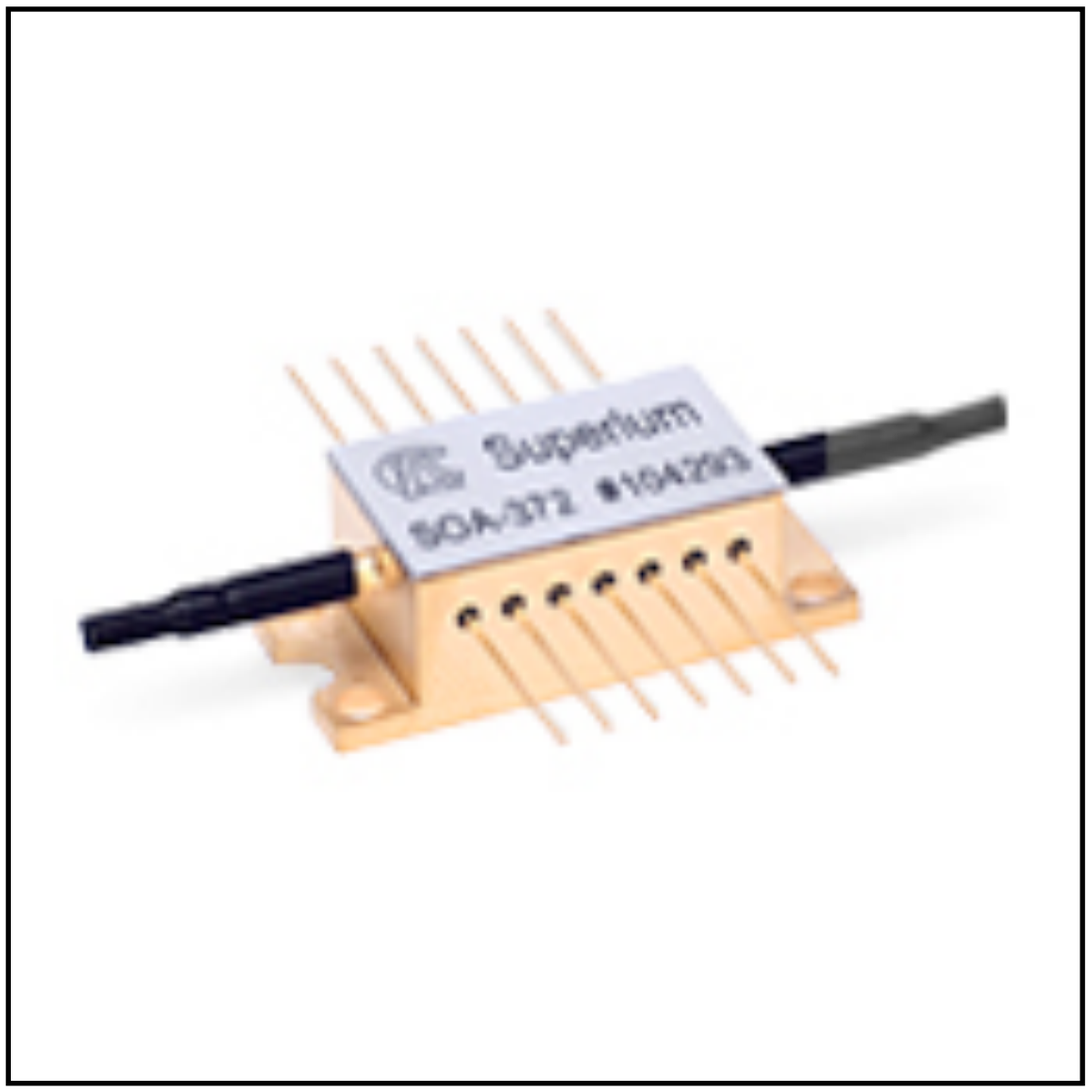
- We had to reach 40km.
- No appliances.
  - Started to work with vendors and design bureaus to get that changed.
- 40km optics might be around the corner.
  - Q4 2012... from one vendor...
  - Electrical power budget might be a problem.

Agenda

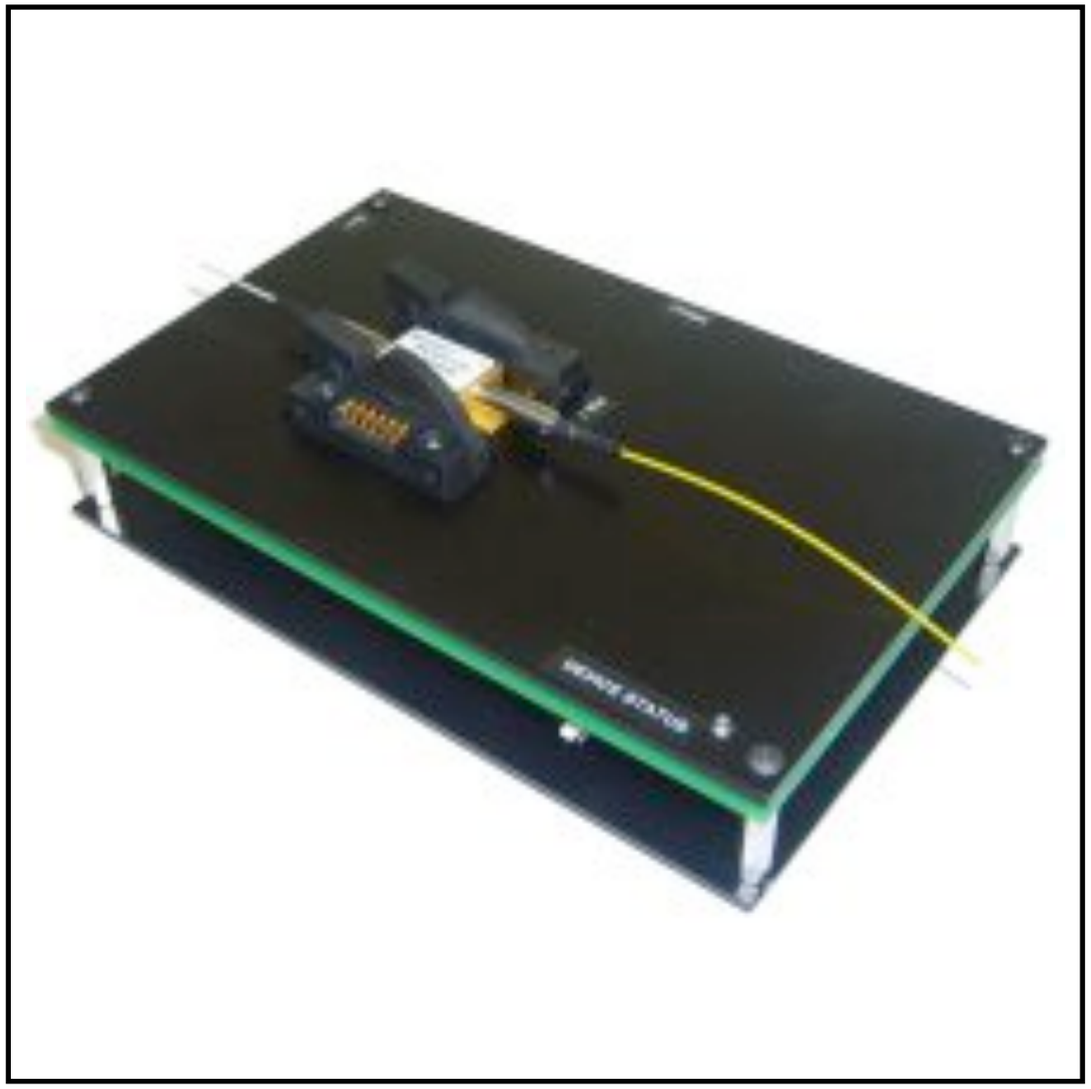
- AMS-IX
- AMS-IX
- 100Gbit/s technology
- Problem statement
- Optical Amplifier development
- Metro DWDM equipment
-

SOA appliance

- We only could find components

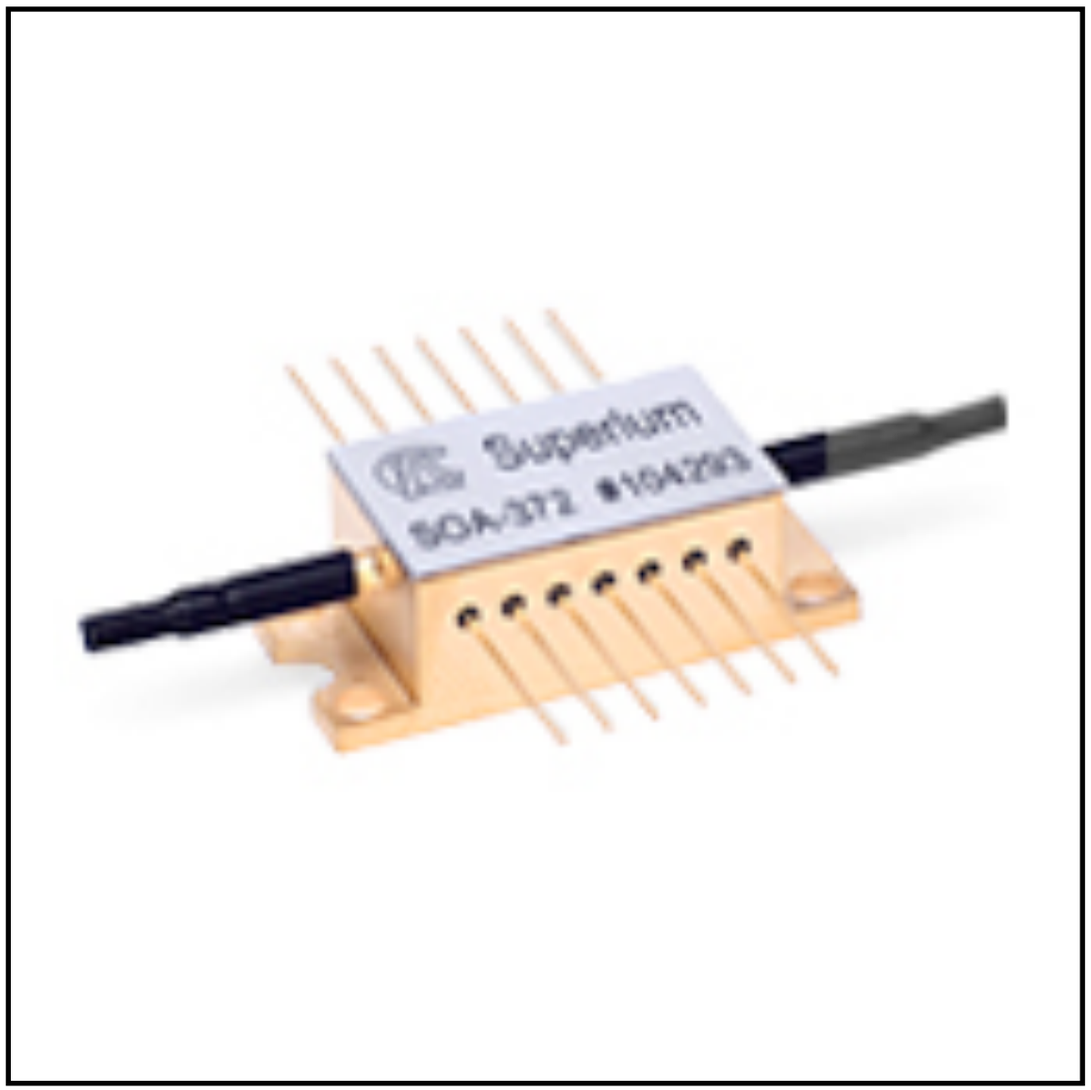


SOA appliance



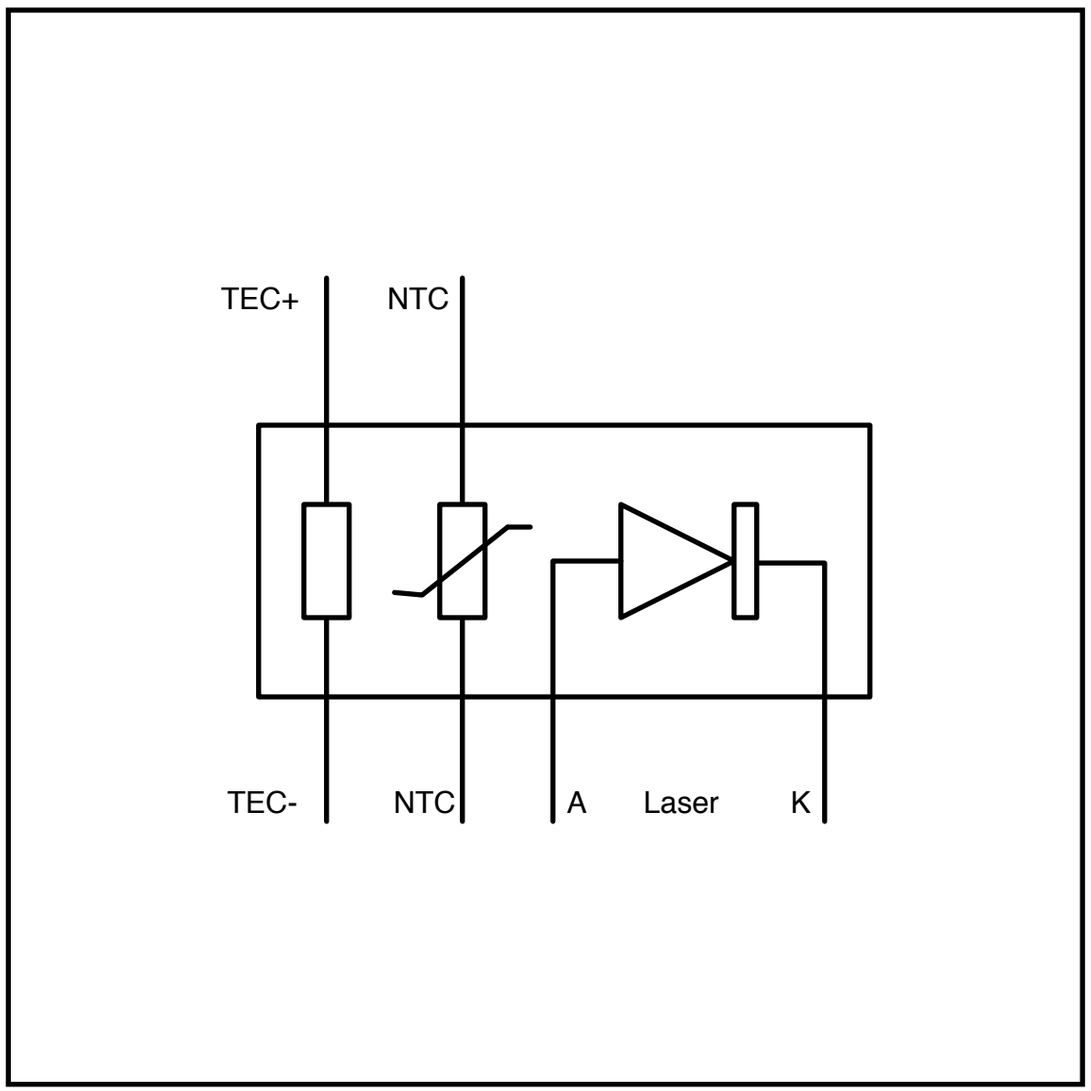
- We only could find components or lab equipment.

SOA appliance



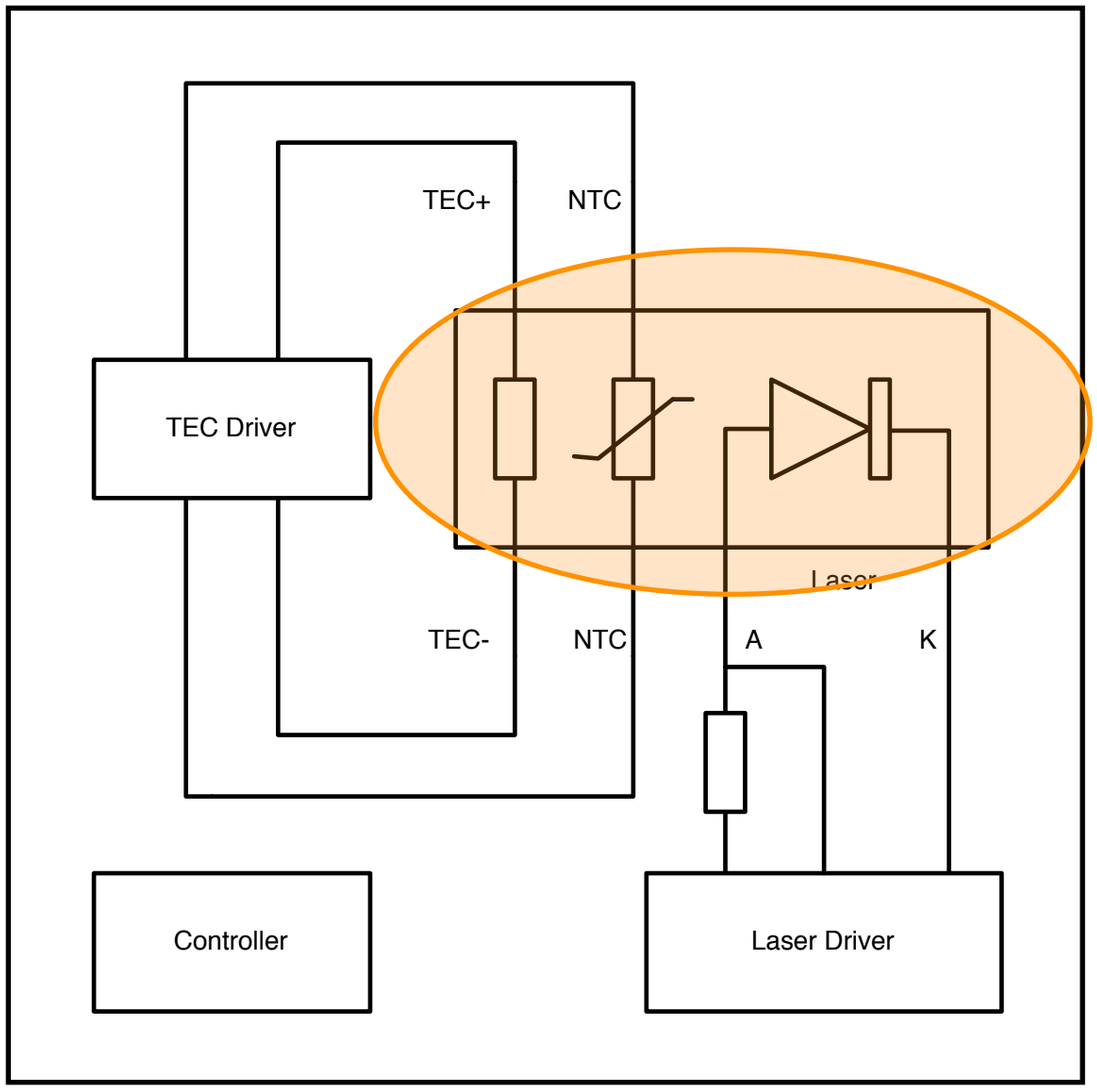
- We only could find components or lab equipment.
- What is in that component?

SOA appliance



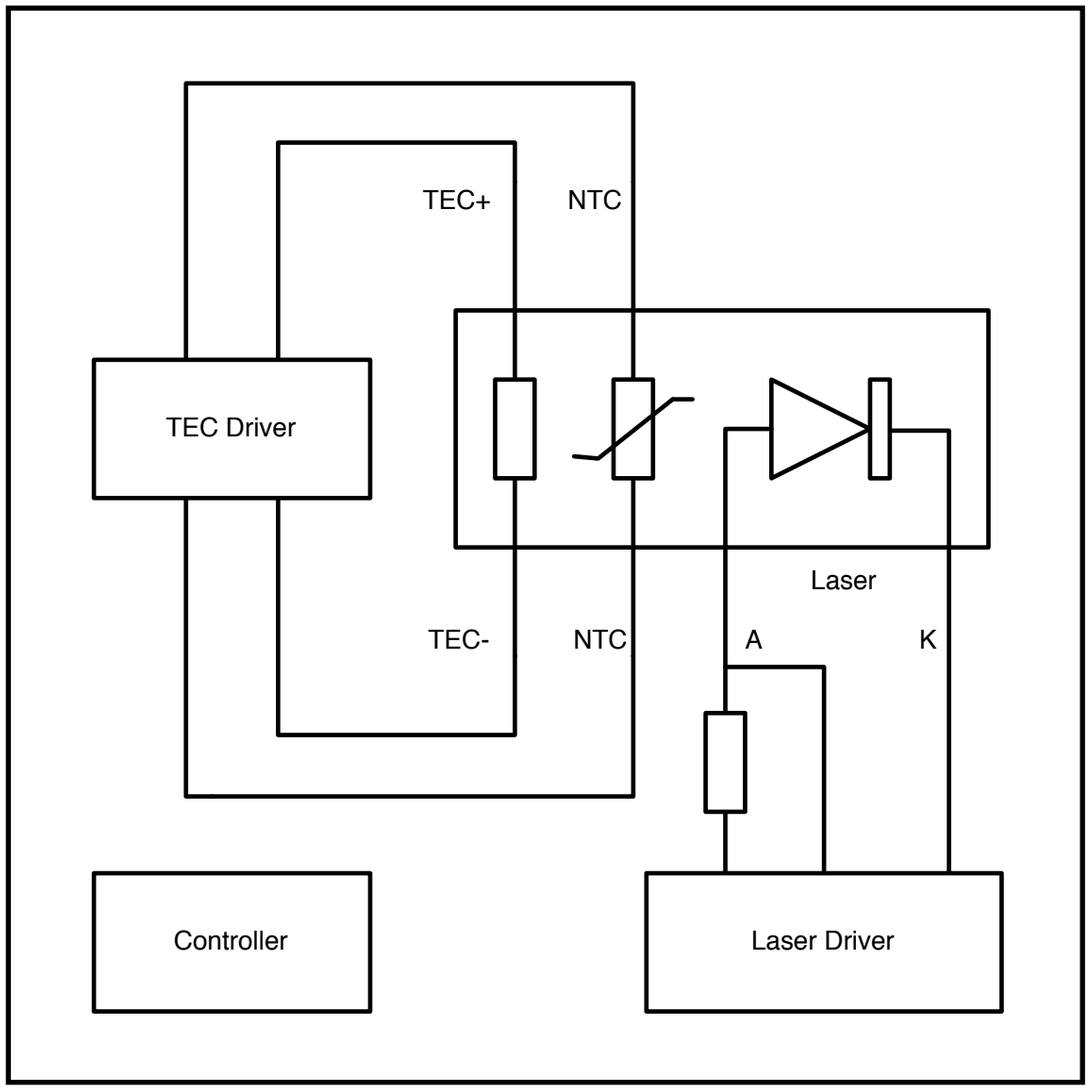
- We only could find components or lab equipment.
- What is in that component?
- SOA consist of three parts.
  - Laser
  - Peltier element (TEC)
  - NTC (temperature sensor)

SOA appliance



- SOA consist of three parts
  - Laser
  - NTC (temperature sensor)
  - Peltier element (TEC)

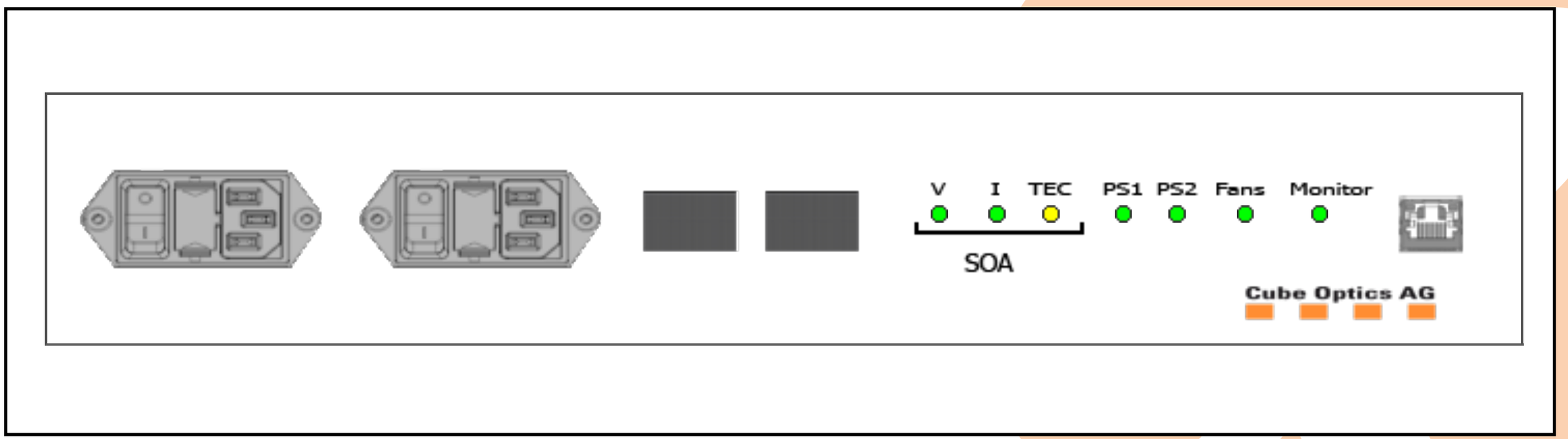
SOA appliance



- SOA consist of three parts
  - Laser
  - NTC (temperature sensor)
  - Peltier element (TEC)
- SOA appliance
  - TEC driver to keep the temperature constant.
    - Feedback loop.
  - Laser driver to control the current through the laser.
    - Feedback loop.
  - Controller

SOA appliances

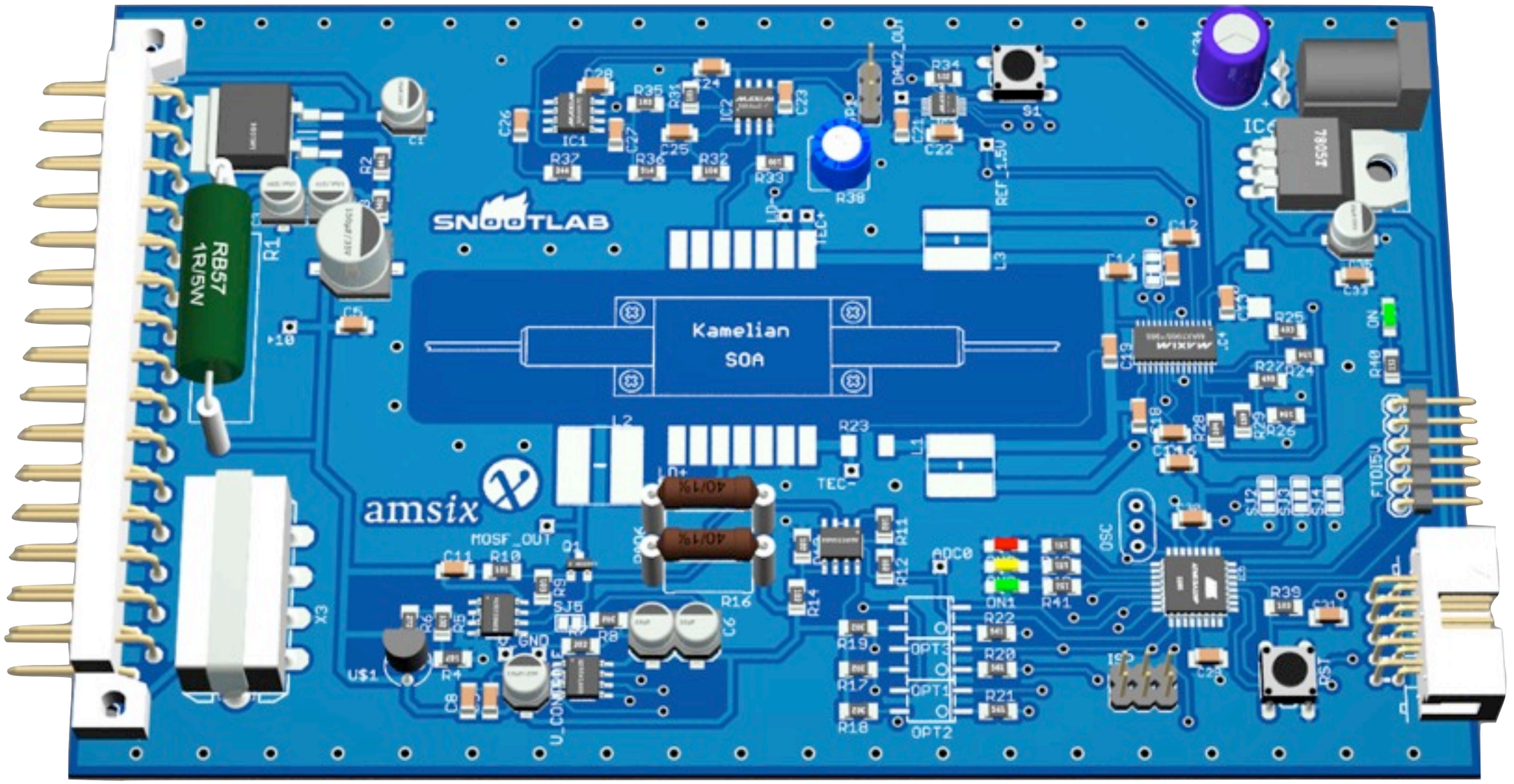
- Cube optics built an SOA appliance
  - 1U 19” rack mountable appliance
  - Dual power supply
  - Web interface
- Available today (with a delivery time)



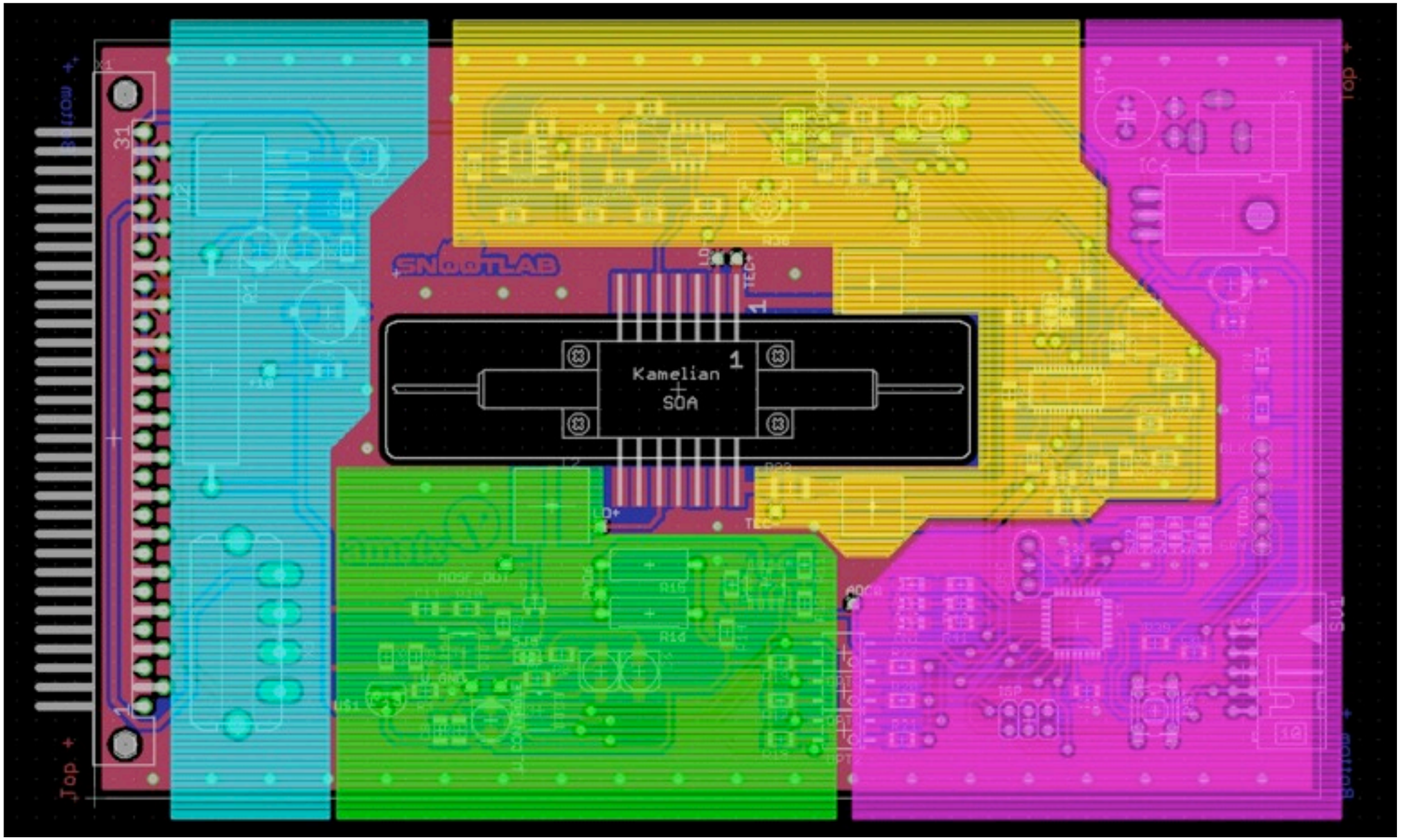
SOA appliances

- SnootLab
  - 3U for a rack of 8 SOAs
    - Covega or Amphotonics (Kamelian)
  - Dual power supply
  - SNMP interface
  - Open Hardware
  - contact@snootlab.com

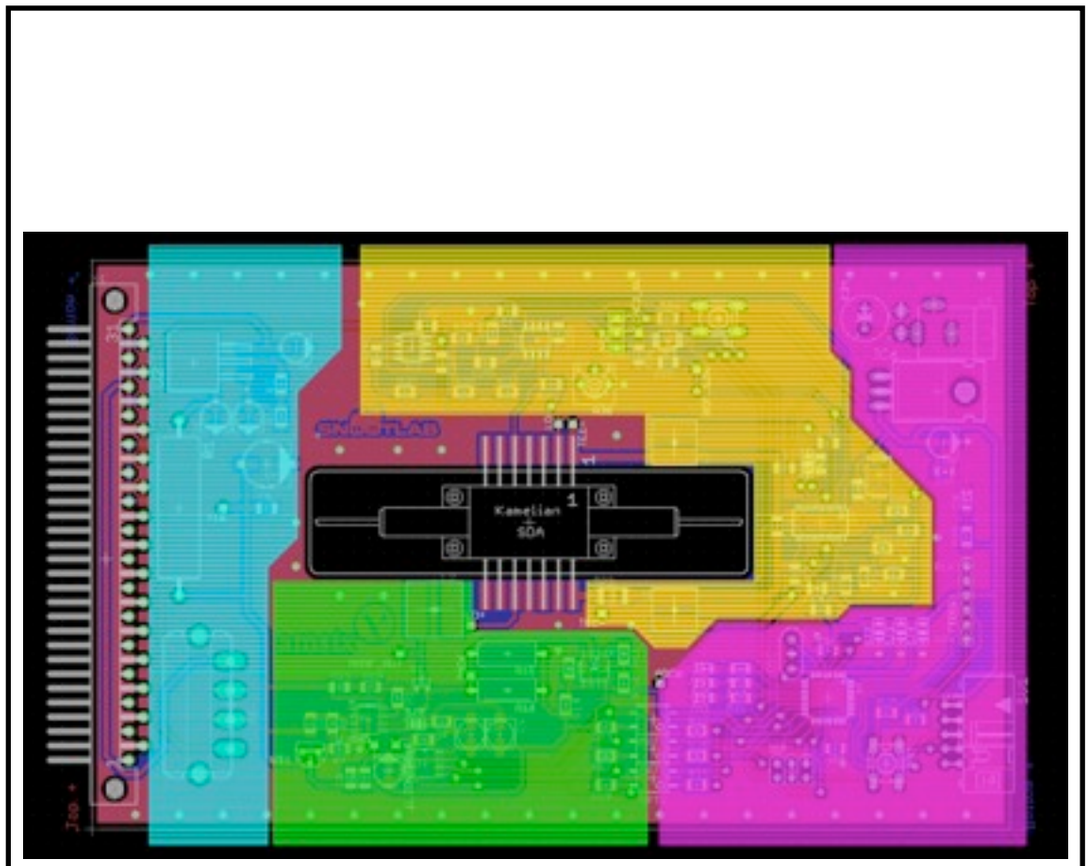
1st Prototype



2nd Design



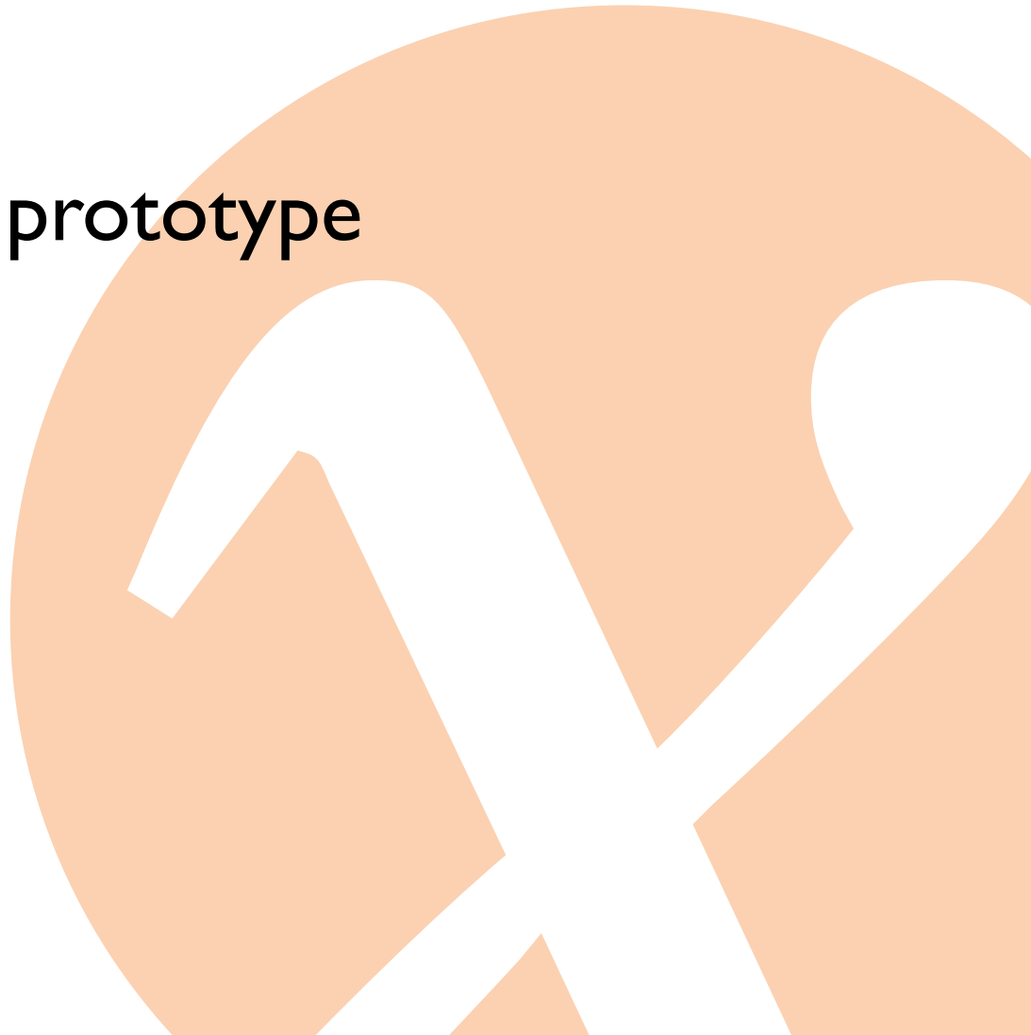
2nd Design



- Blue the power management.
- Yellow : the TEC (Peltier) management.
- Green: the Laser driver.
- Pink : the "arduino".
  - Controller.
  - The LED power management.
- Multi vendor:
  - Amphotonics (Kamelian).
  - ThorLabs (Covega).

Test results

- Error free over 22 km over 96 hours.
  - Cube optics
- Not error free over 45 km.
  - Both Cube optics and SnootLab's prototype
  - Might be too long
  - Might be wrong settings
    - More experiments are needed



Test results

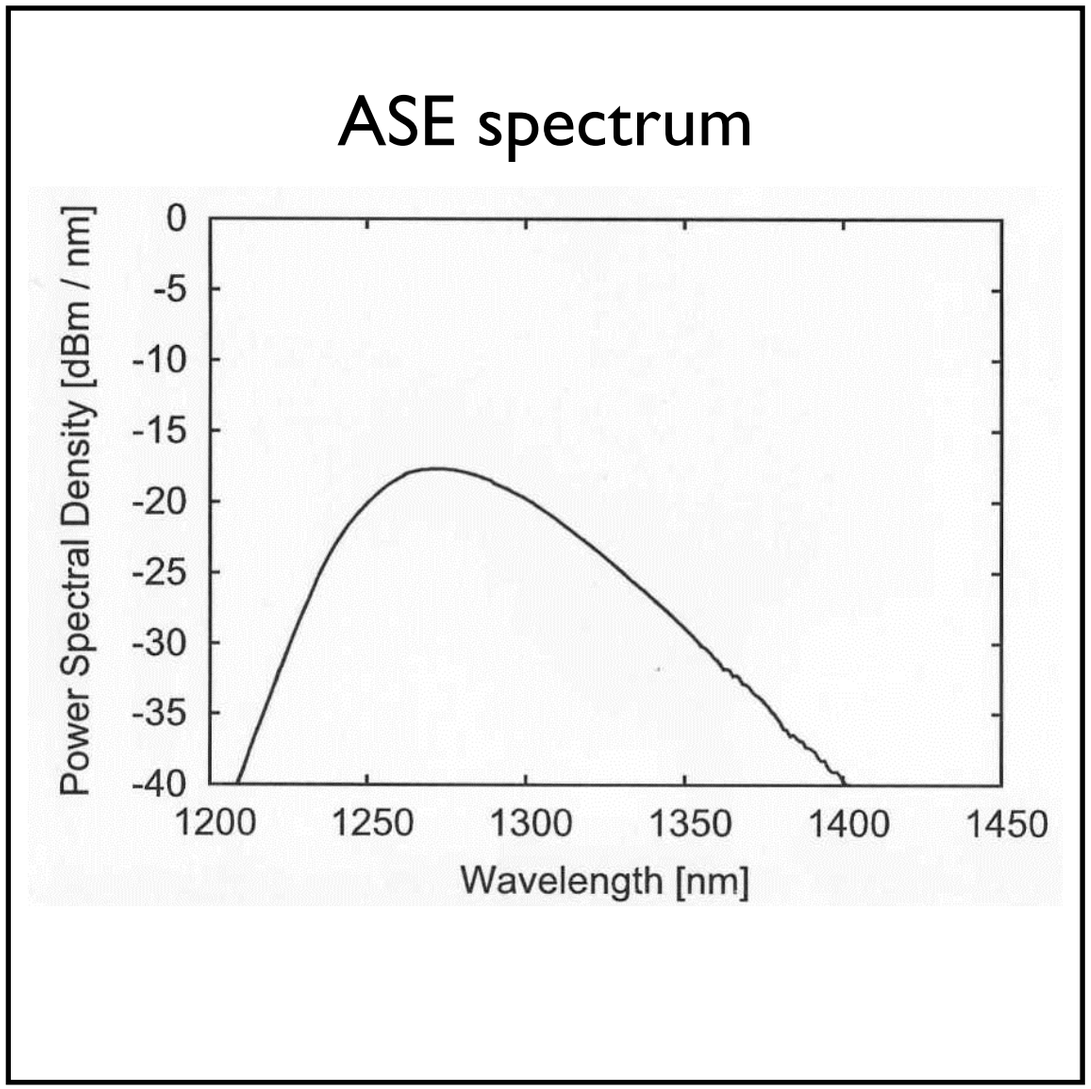
- Cube optics
    - Error free over 22.6 km over 96 hours.
- SnootLabs
    - Error free over 22.6 km over 36 hours.
        - Started the tests monday night.
- This is suitable for most of our needs.

Test results

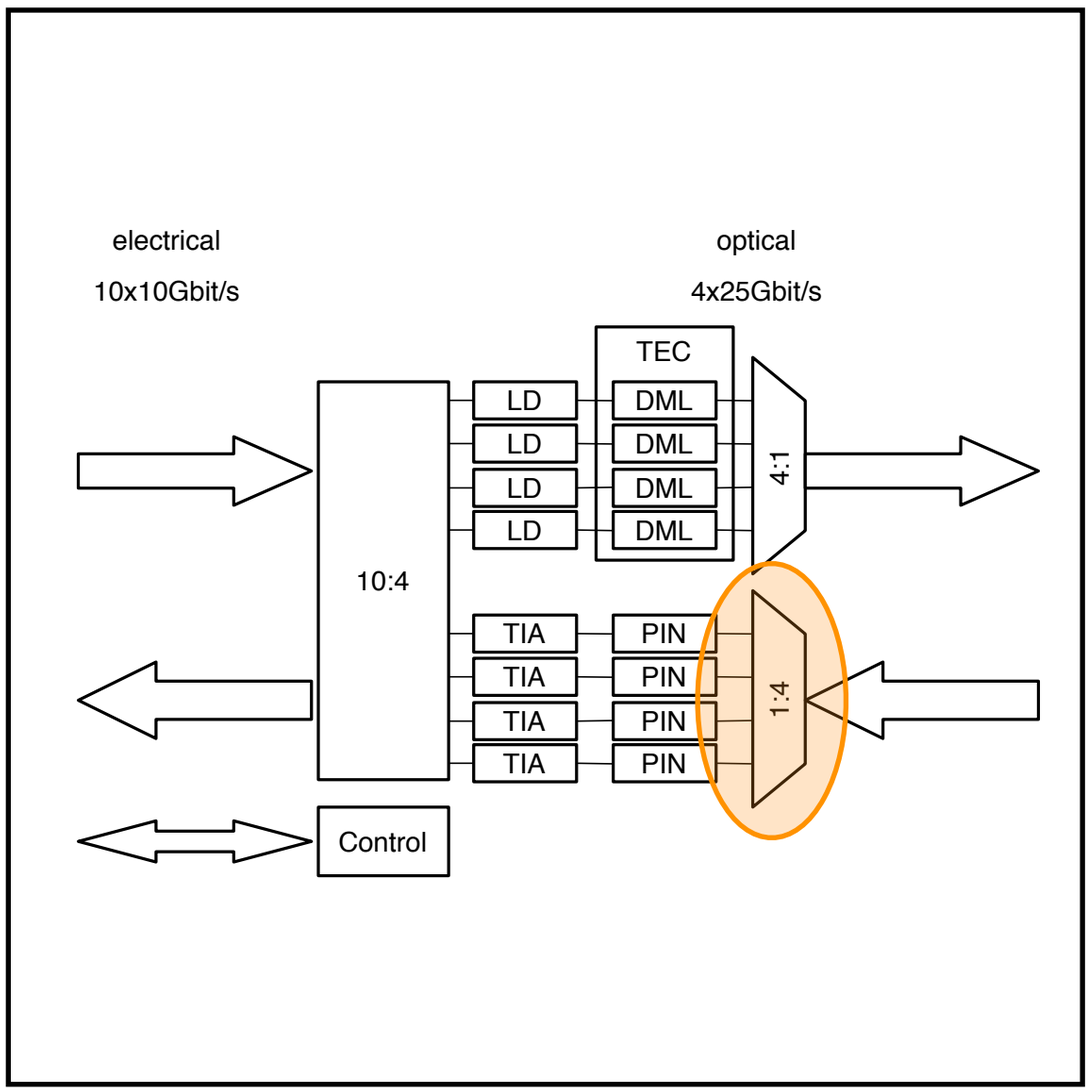
- Not error free over 45 km.
  - Both Cube optics and SnootLab's 1st prototype.
  - Might be too long.
    - Dispersion?
  - Extinction ratio
    - Ratio between 0 and 1 levels.
    - Amplifier might not amplify these levels equally.
- Noise?

Noise

- Amplified Spontaneous Emission (ASE)

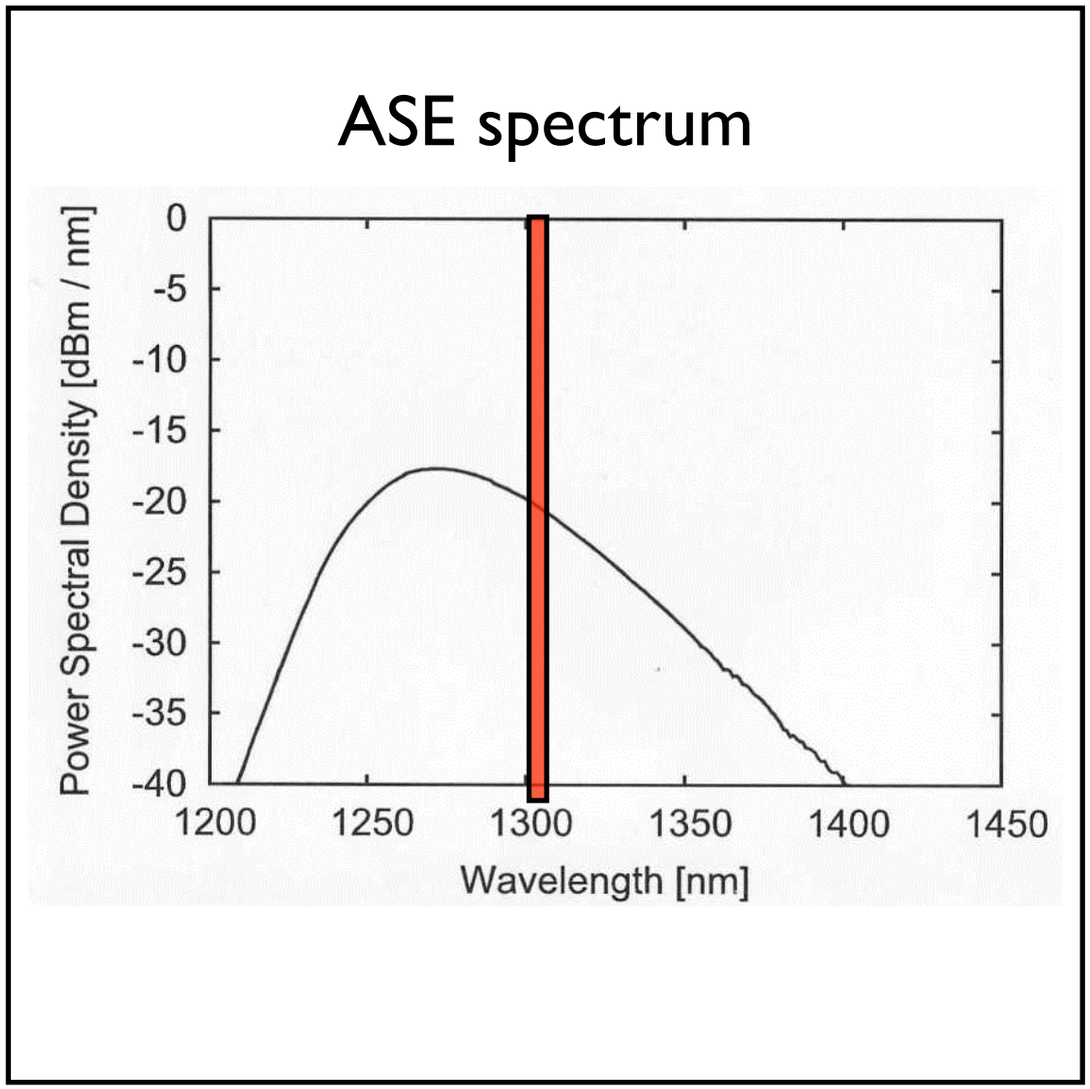


Noise



- 100G-LR receivers
- LAN WDM filter in the receiver drops most of the noise spectrum.

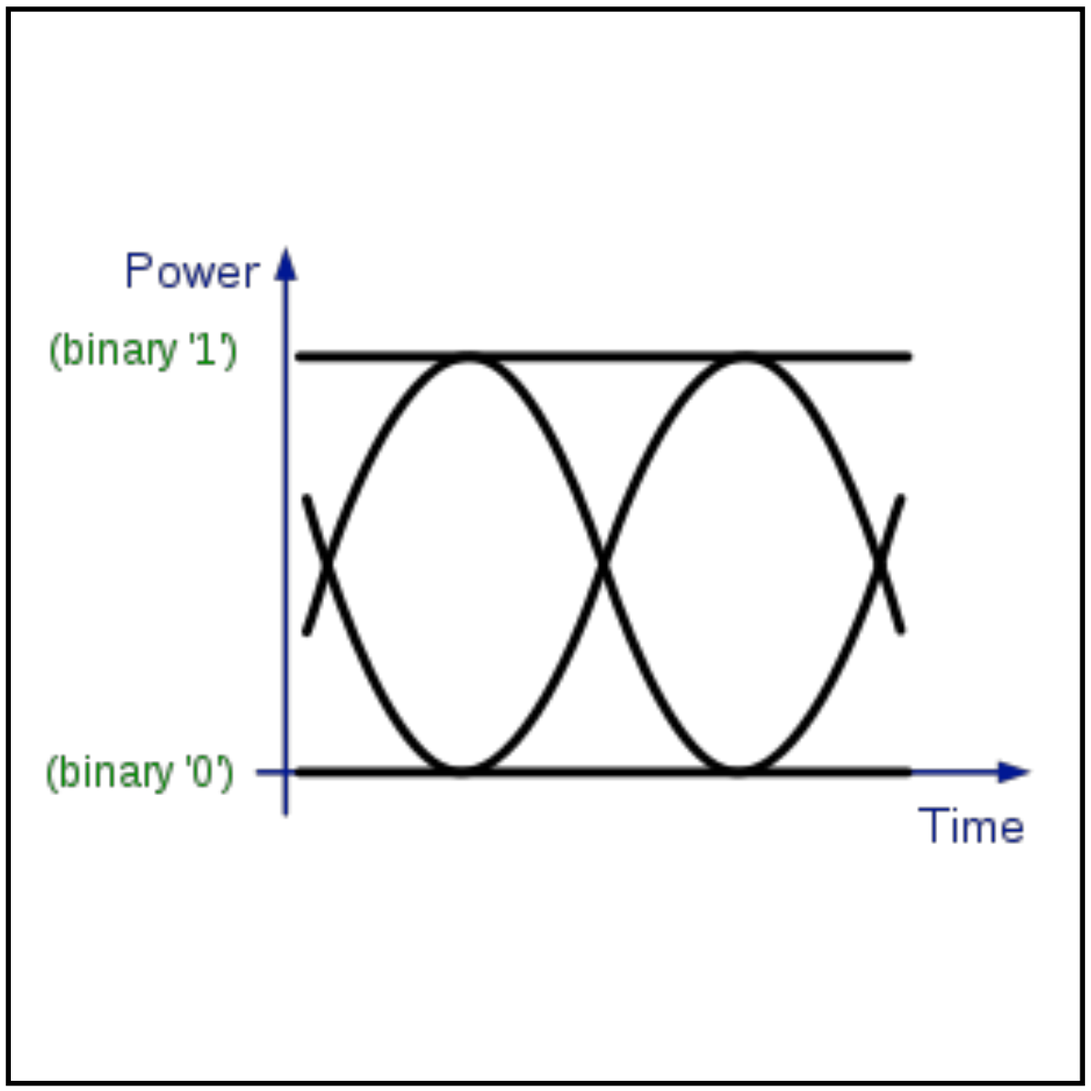
Noise



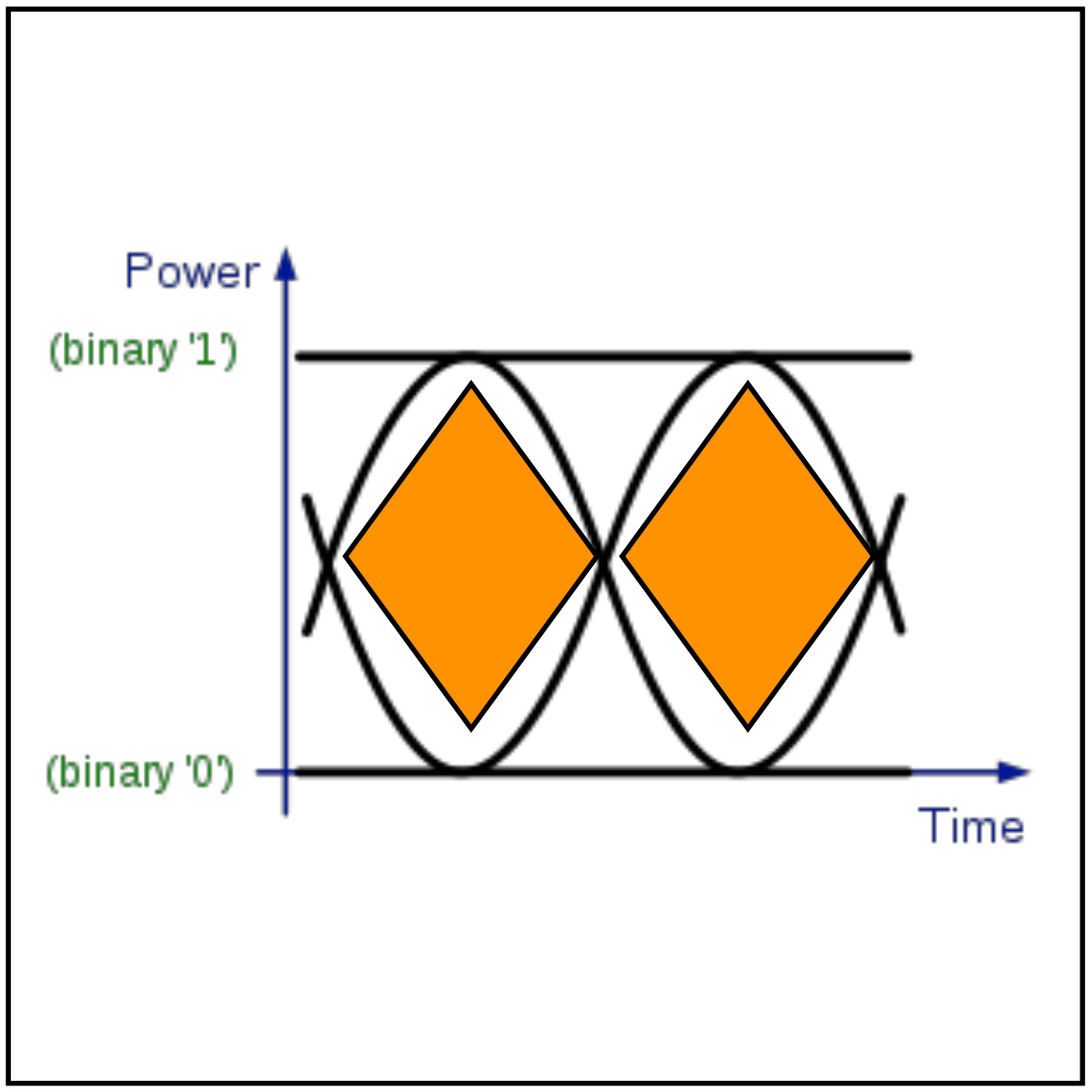
- Amplified Spontaneous Emission (ASE)
- LAN WDM filter in the receiver drops most of the noise spectrum.
- Noise should be suppressed by 40dB to the signal.

Extinction ratio

- Difference between 0 and 1 level.

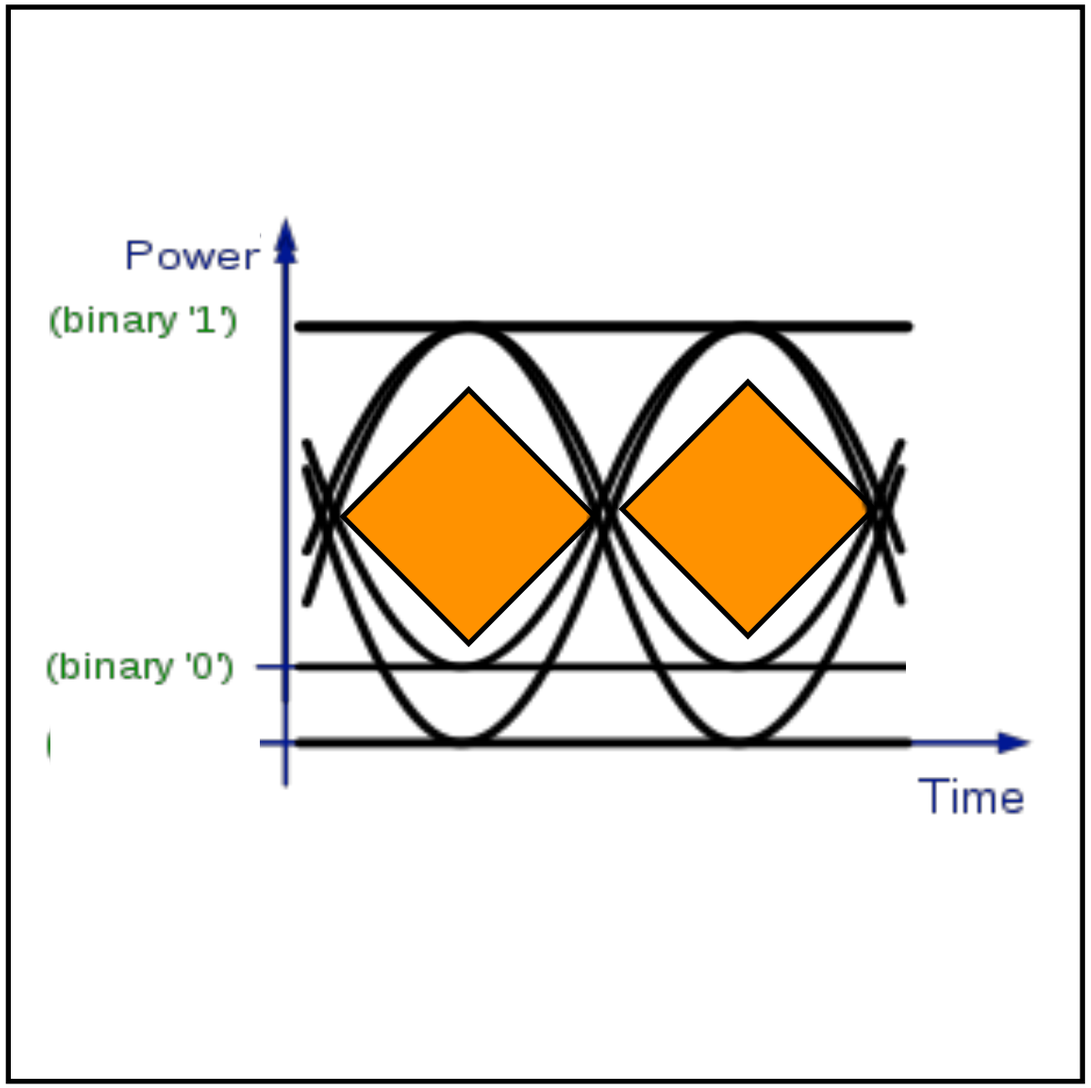


Extinction ratio



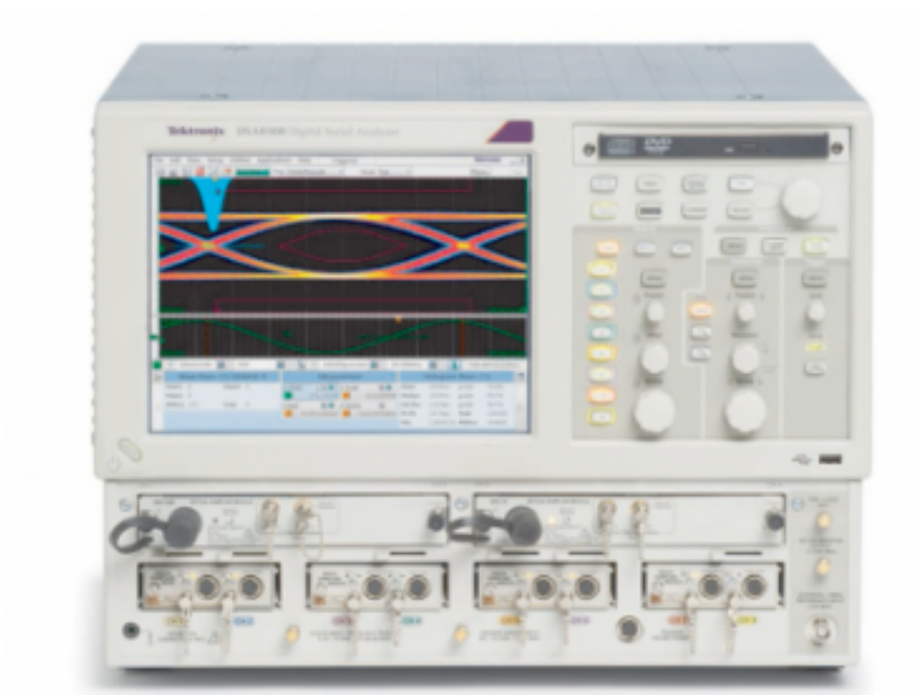
- Difference between 0 and 1 level.
- Eye pattern

Extinction ratio



- Difference between 0 and 1 level.
- Eye pattern
- 0 Level might be amplified more than the 1 level.
- Unlikely considering the low currents.

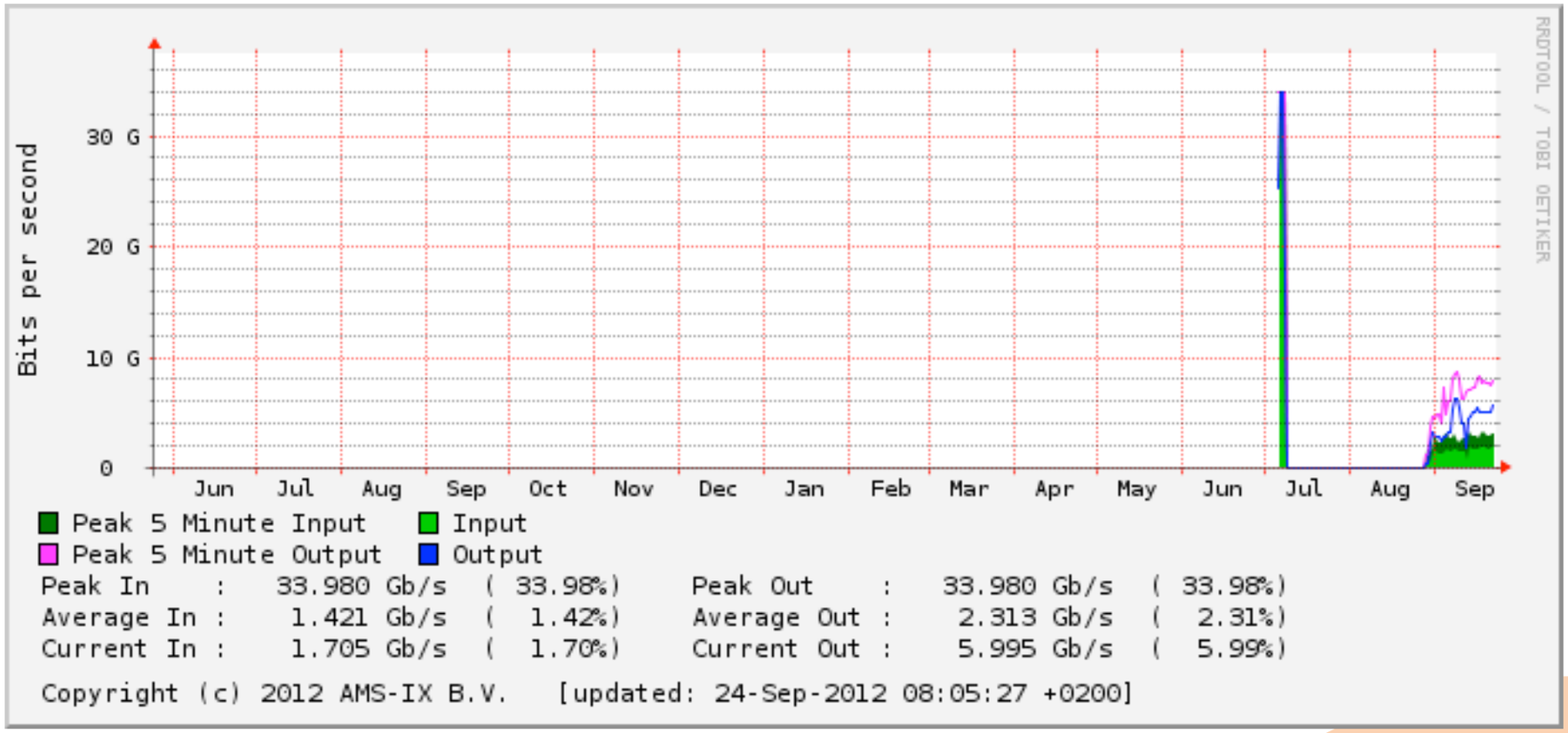
Extinction ratio

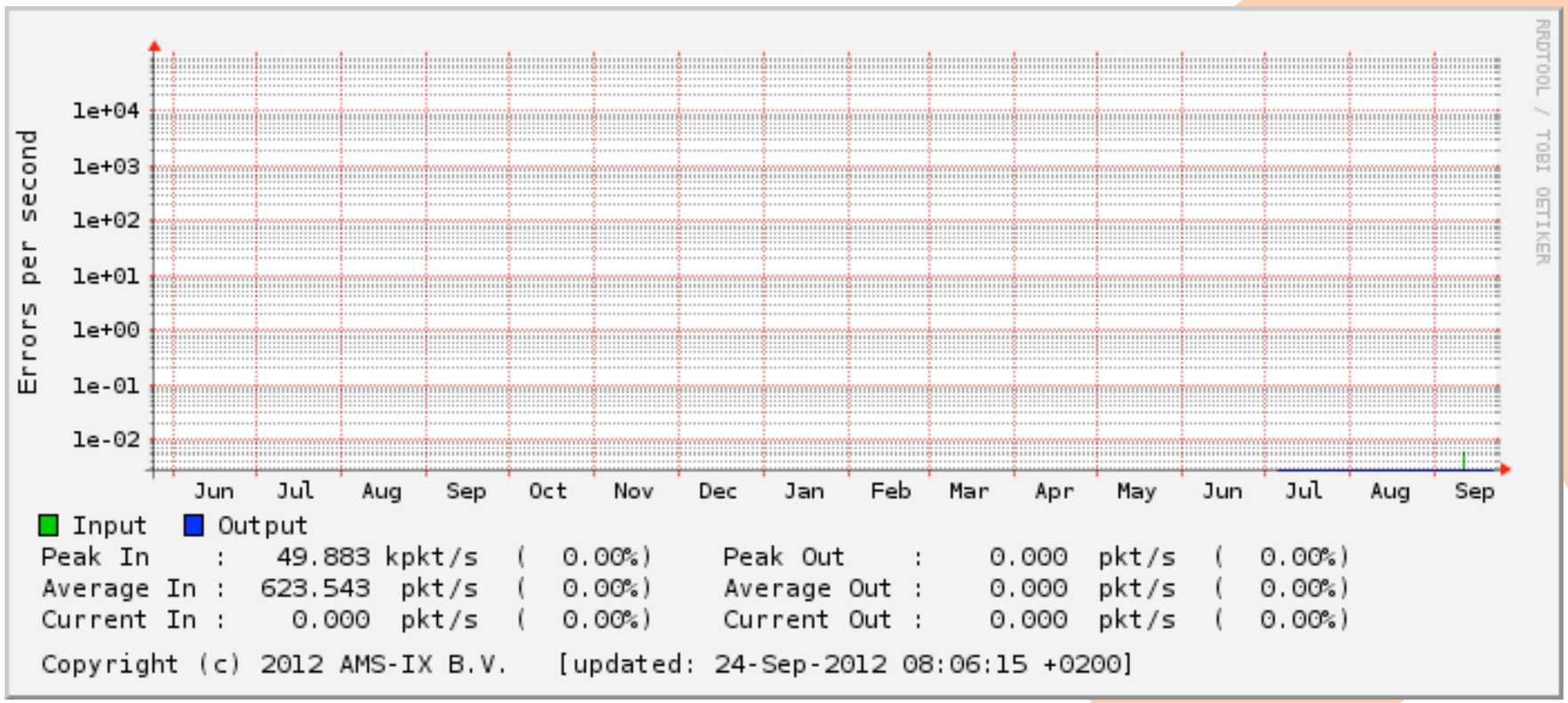


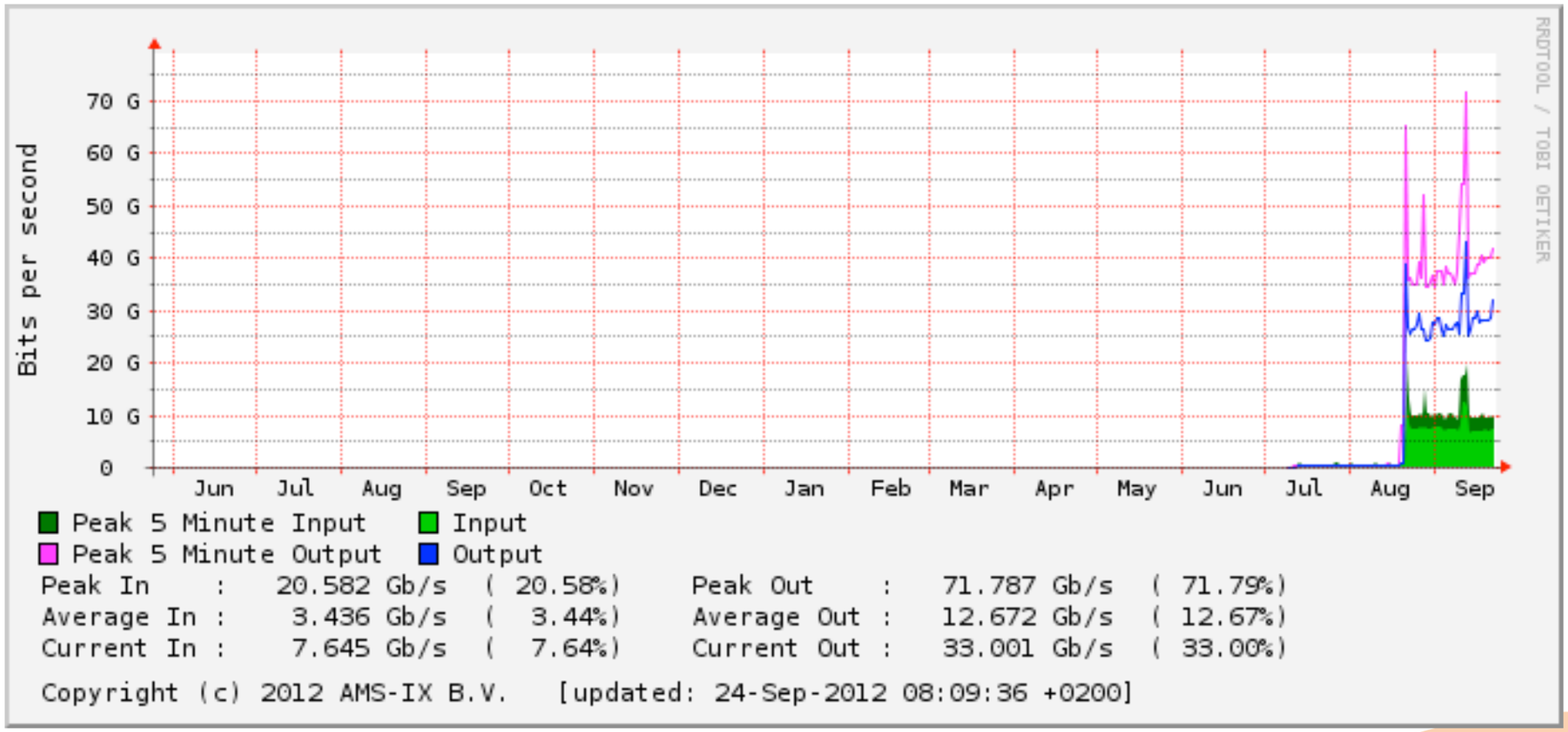
- Difference between 0 and 1 level.
- Eye pattern
- 0 Level might be amplified more than the 1 level.
  - Unlikely considering the low currents.
- We don’t have the equipment to see this.
  - Tektronix DSA8300
  - 80C10B module

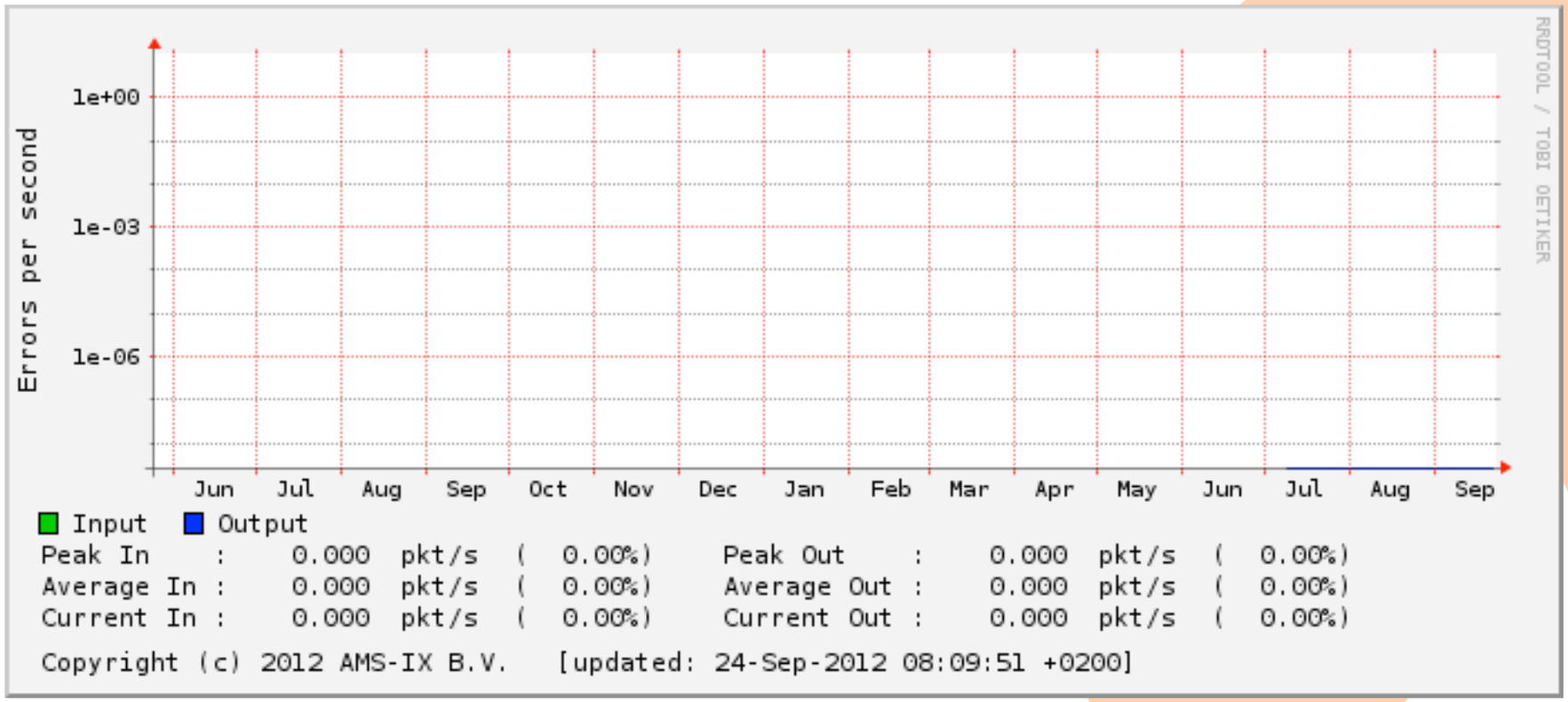
Conclusion

- SOA and PDFA
  - Certainly suitable for ~ 25km ranges
  - Errors on 45km
    - Not sure where they come from yet.
    - Transmission equipment from ADVA, MRV, etc


RRDTOOL / TOBI OETIKER
Bits per second
30 G
20 G
10 G
0
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep
Peak 5 Minute Input
Input
Peak 5 Minute Output
Output
Peak In : 33.980 Gb/s (33.98%)
Average In : 1.421 Gb/s (1.42%)
Current In : 1.705 Gb/s (1.70%)
Peak Out : 33.980 Gb/s (33.98%)
Average Out : 2.313 Gb/s (2.31%)
Current Out : 5.995 Gb/s (5.99%)
Copyright (c) 2012 AMS-IX B.V. [updated: 24-Sep-2012 08:05:27 +0200]


RRDTOOL / TOBI OETIKER
Errors per second
1e+04
1e+03
1e+02
1e+01
1e+00
1e-01
1e-02
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep
Input
Output
Peak In : 49.883 kpkt/s (0.00%)
Average In : 623.543 pkt/s (0.00%)
Current In : 0.000 pkt/s (0.00%)
Peak Out : 0.000 pkt/s (0.00%)
Average Out : 0.000 pkt/s (0.00%)
Current Out : 0.000 pkt/s (0.00%)
Copyright (c) 2012 AMS-IX B.V. [updated: 24-Sep-2012 08:06:15 +0200]


Bits per second
70 G
60 G
50 G
40 G
30 G
20 G
10 G
0
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep
RRDTOOL / TOBI OETIKER
Peak 5 Minute Input
Input
Peak 5 Minute Output
Output
Peak In : 20.582 Gb/s (20.58%)
Average In : 3.436 Gb/s (3.44%)
Current In : 7.645 Gb/s (7.64%)
Peak Out : 71.787 Gb/s (71.79%)
Average Out : 12.672 Gb/s (12.67%)
Current Out : 33.001 Gb/s (33.00%)
Copyright (c) 2012 AMS-IX B.V. [updated: 24-Sep-2012 08:09:36 +0200]


Errors per second
1e+00
1e-03
1e-06
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep
RRDTOOL / TOBI OETIKER
Input
Output
Peak In : 0.000 pkt/s (0.00%)
Average In : 0.000 pkt/s (0.00%)
Current In : 0.000 pkt/s (0.00%)
Peak Out : 0.000 pkt/s (0.00%)
Average Out : 0.000 pkt/s (0.00%)
Current Out : 0.000 pkt/s (0.00%)
Copyright (c) 2012 AMS-IX B.V. [updated: 24-Sep-2012 08:09:51 +0200]

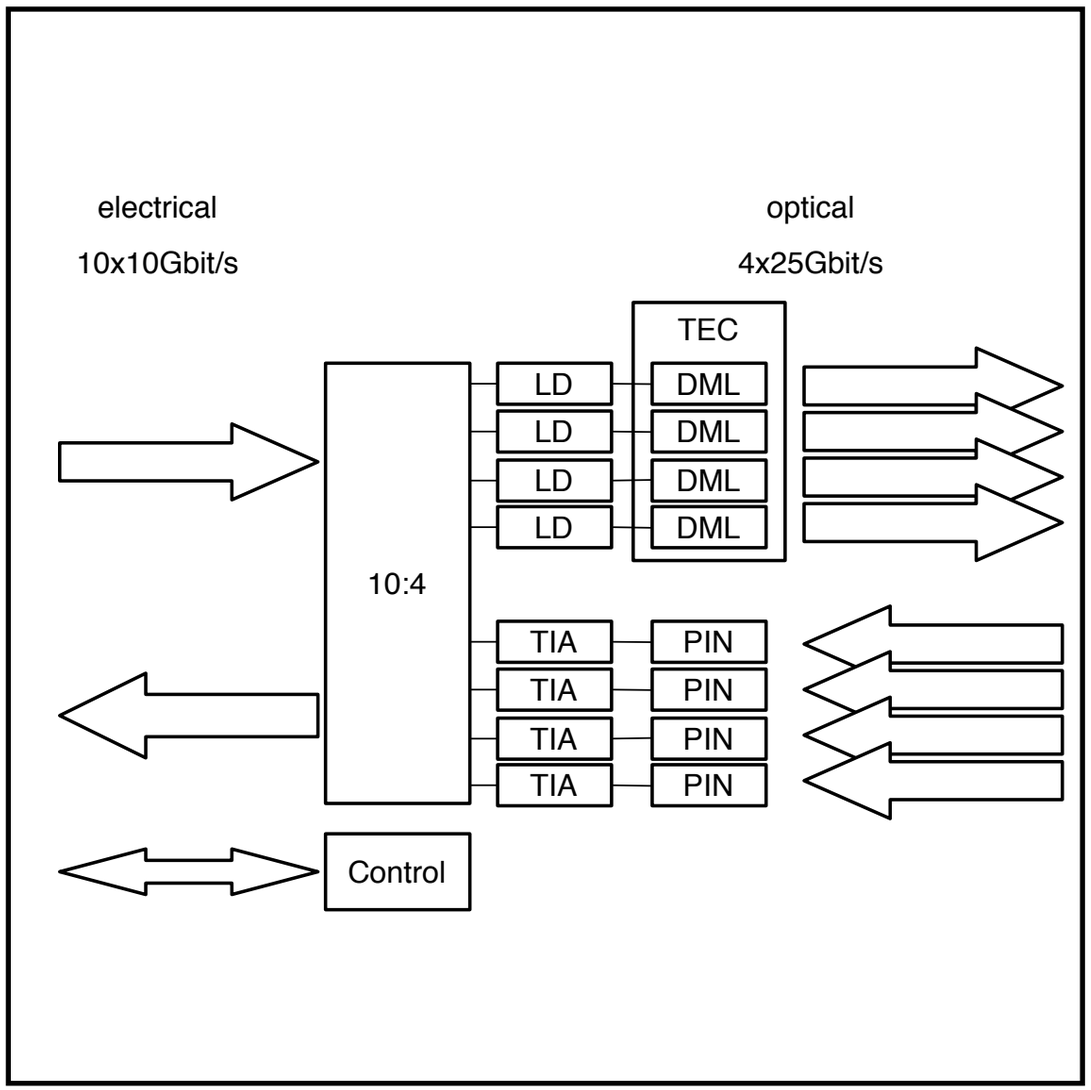
Production results

- SOA
  - 32km and 22km.
  - Errors while tuning in.
    - No plug and play technology.
- No to a few errors while in production.

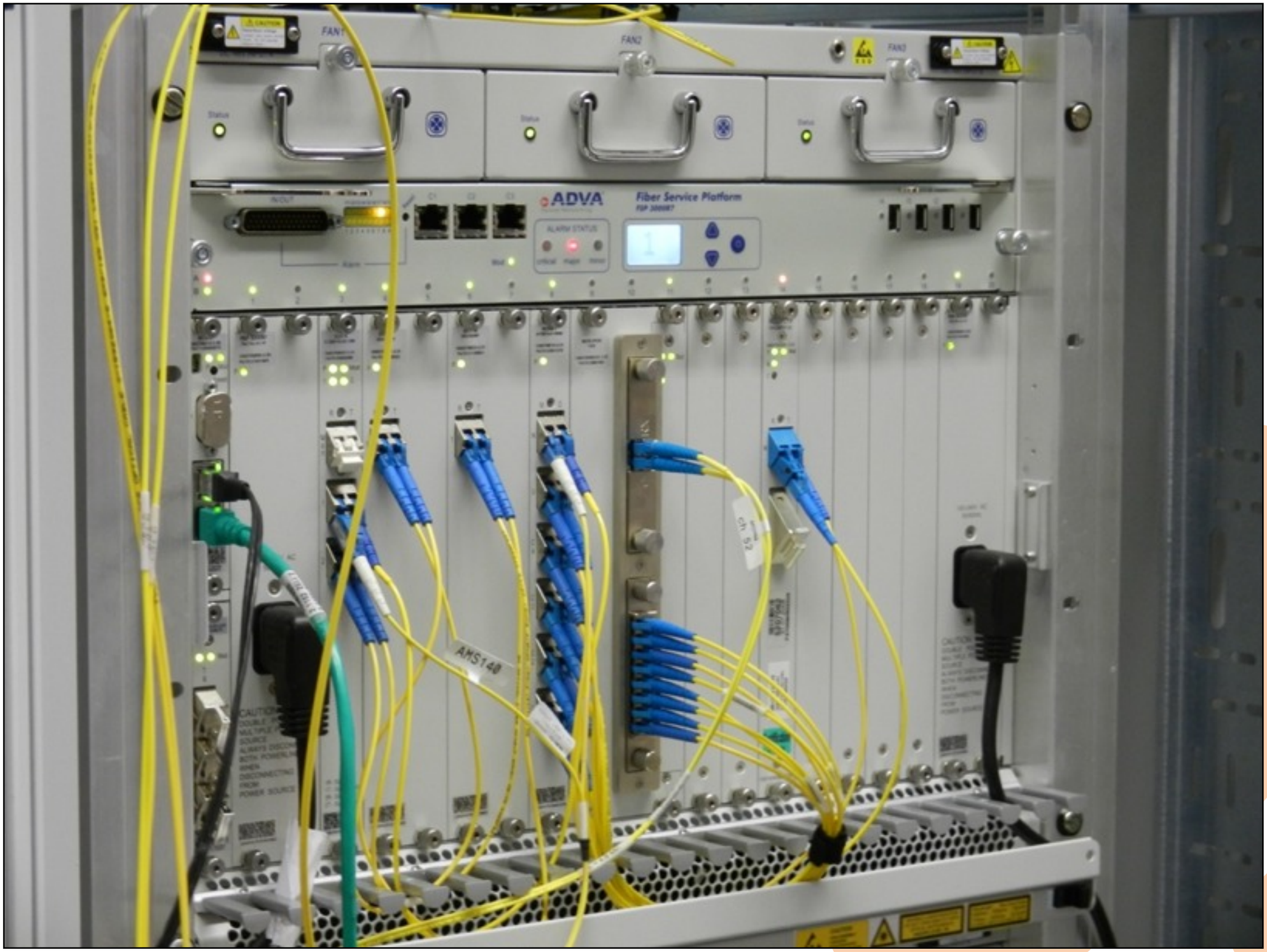
Agenda

- AMS-IX
- 100Gbit/s technology
- Problem statement
- Optical Amplification
- Metro DWDM equipment

Metro DWDM



- We also installed ADVA equipment
  - Finisar CFP without built in muxes
  - 1550nm 4 x 28Gbit/s on 50GHz ITU grid
  - FEC
  - Successful over 45km where SOA could not make it error free.


FAN1
FAN2
FAN3
ADVA
Fiber Service Platform
AMS140

End

Comments & Questions