Scaling your metro
- the journey from 10G to 400G

Will Hargrave

Virtual UKNOF

20th July 2020 | Everywhere

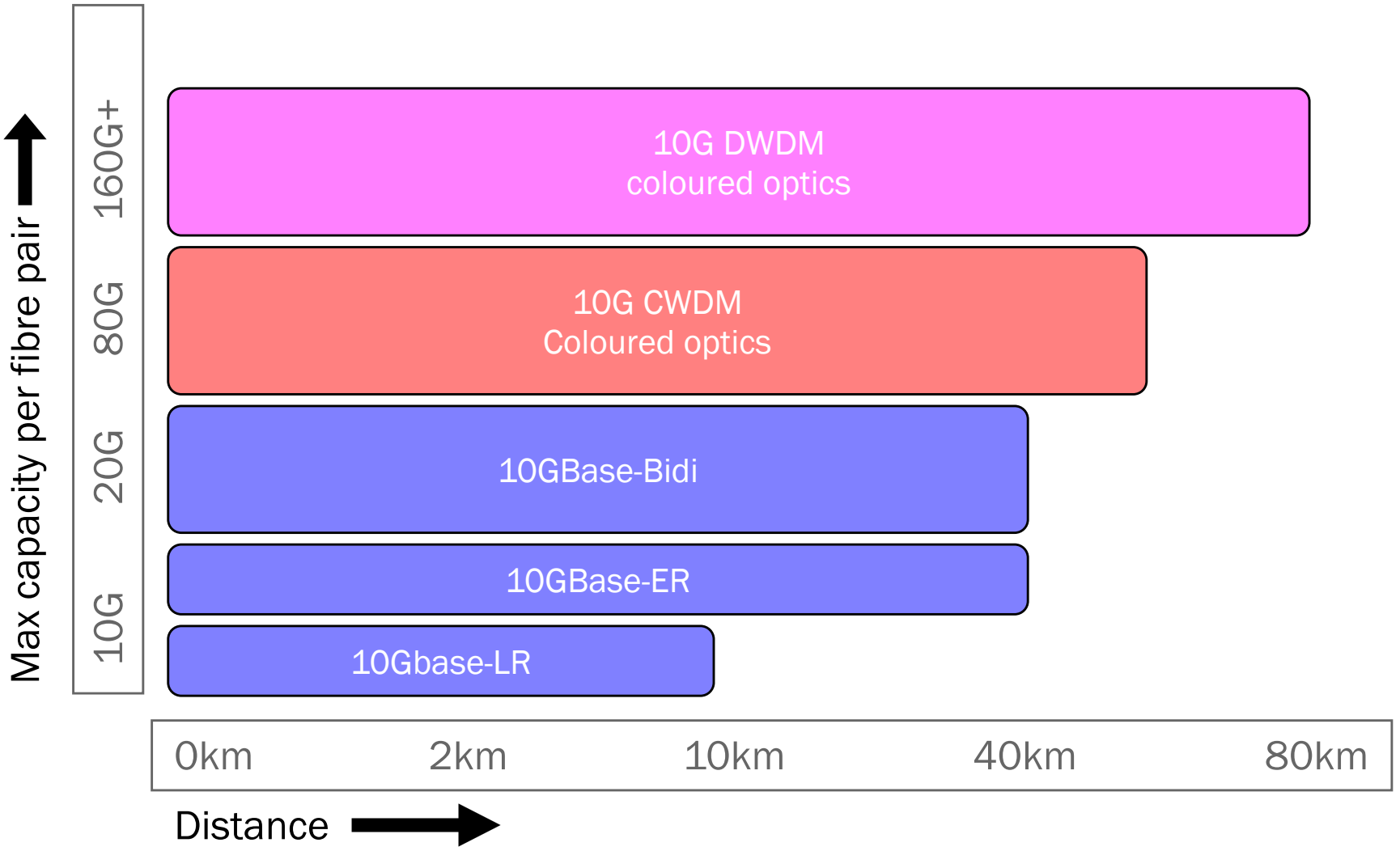
Typical Internet Exchanges

- Ethernet fabric shifting large volumes of traffic in the metro
  - What is ”metro”? 10km? 80km? 200km?
  - London: longest path between interconnection dense colos is 72km
  - Many links are under 10km
- Technology choices impact viability
- Typically point to point links on dark fibre to enable scaling
  - Solutions ideally scale at less-than-linear
- We will focus on technologies suitable for Internet Exchanges
  - Own infrastructure, scaling, point-to-point, low cost

ISL Sizing Strategy at LONAP

- ISLs must deliver traffic demand, even during fault conditions
- We use sflow to report intersite traffic demand
- Sum of per-site connected capacity is not very useful
  - Wide variation in port utilisations
- Custom graphing solution to sum ”LAG” members
  - All ISLs are multipath ECMP IP p2ps, no LACP
  - Must also check traffic balance within a ‘LAG’
- 30% traffic growth per year

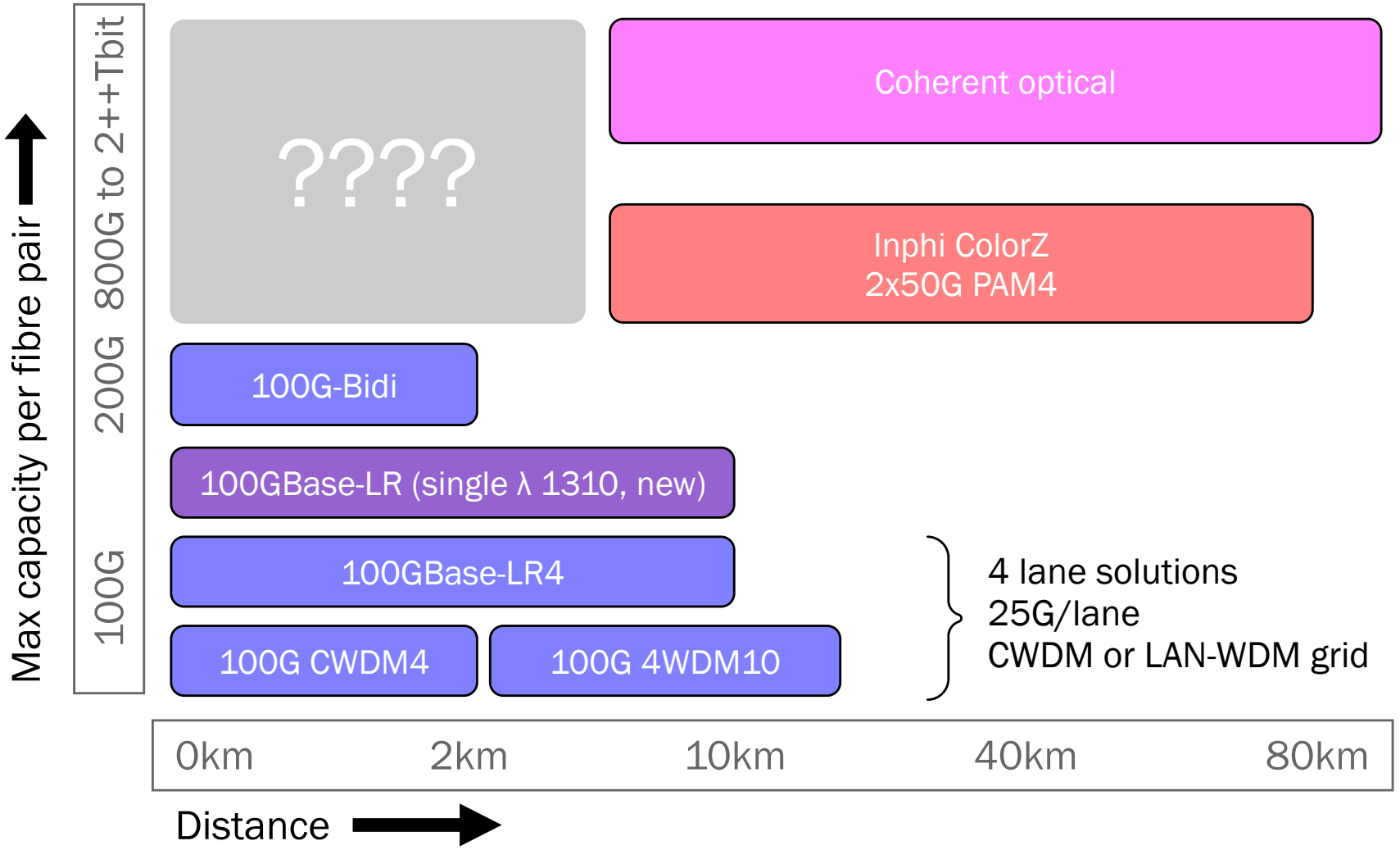
10G era technologies



2015 era – 10G DWDM

- Simpler times
- Buy 40km or 80km DWDM SFP+, insert into switch, passive mux
- No amplification, dispersion compensation to worry about
- n * 10GE – scaled to around 160G in our current setup
- Difficulties with ’too many moving parts’ at this scale
  - Total of 32 SFP+, 64 fibre cores, 128 connections
  - Could be > 256 fibre end-faces to clean even for a modest 16 channel setup

100G era technologies



100G choices at LONAP

2016:

- Lighting 100G on dark fibre pairs/ 1310nm waves
- Initially ”standard” 100GBase-LR4

2018 onwards:

- Migrated to CWDM4/4WDM10 for 30-50% cost saving
- 4WDM10 provides superior optical performance than LR4
- Both require RS-FEC host support (not a problem)

2018-20:

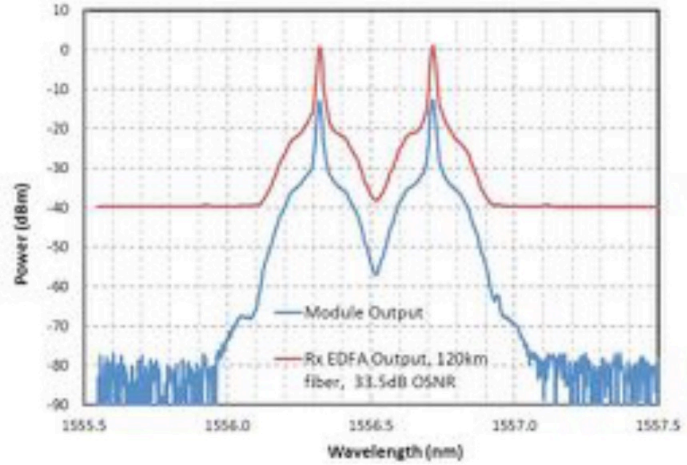
- Need more than 100G per path
- Deployed 100G PAM4 solution – scales to 800G+/pair
- Testing 100G Bidi – scales to 200G/pair

100G+ in the metro (10km+)

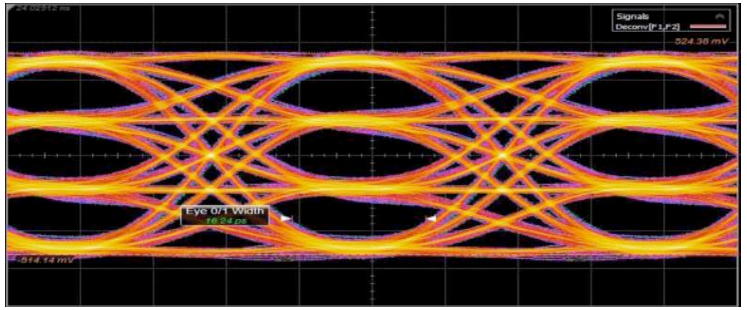
- No straightforward passive 100G DWDM option
- Investigated Coherent transponder solutions - many options exist
- Coherent can deliver up to 600G/wavelength, over thousands of kilometres at many terabits/pair – prime technology for larger telcos
- Flexible modulation
- Relatively power-hungry and complex
- Relatively poor financial scaling
- *Opinion: "I don't feel that these are a great solution for IXPs, is there something cheaper/better at scaling?"*

100G PAM4

- Inphi ”ColorZ”: 2x50G PAM4 λ encoded in a single 100GHz DWDM channel
- Standard QSFP28 in switch
  - Arista EOS supports DOM extensions to instrument on-optic FEC DSP
- Optical specifications:
  - Output power of -8dBm and min input power of -2dBm
  - Dispersion tolerance +/-6 km on G.652 fibre
  - Requires dedicated optical line system to address these challenges



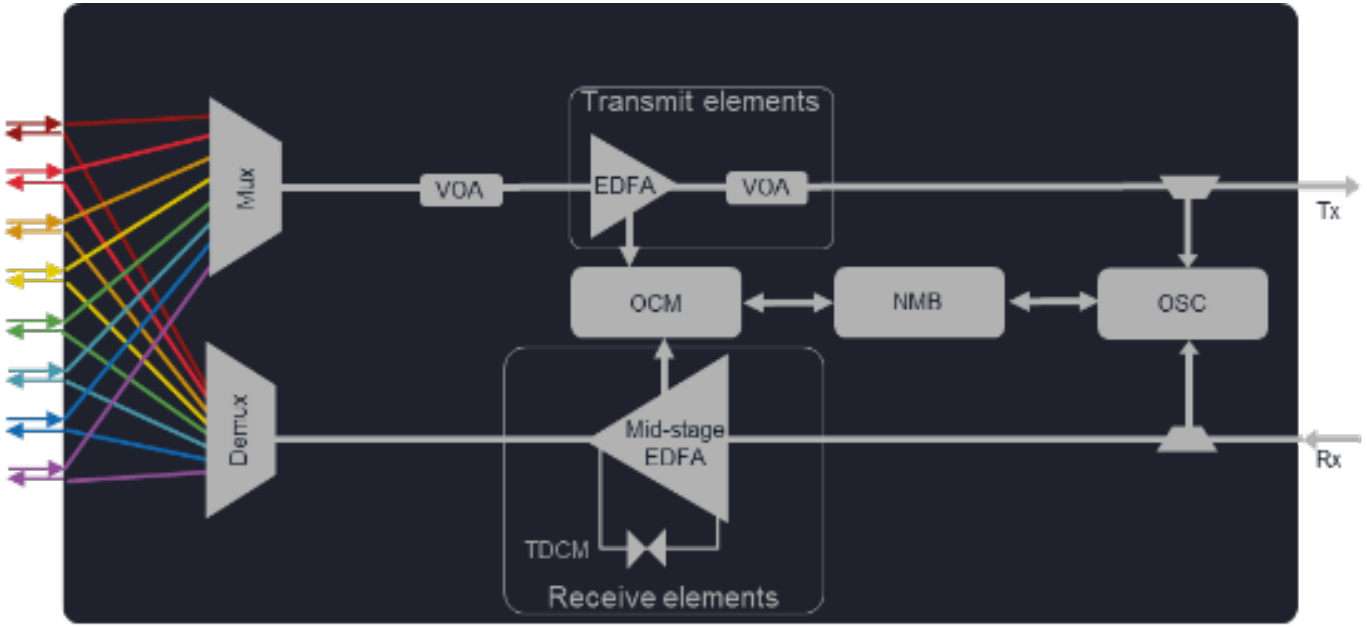
PAM4 modulation



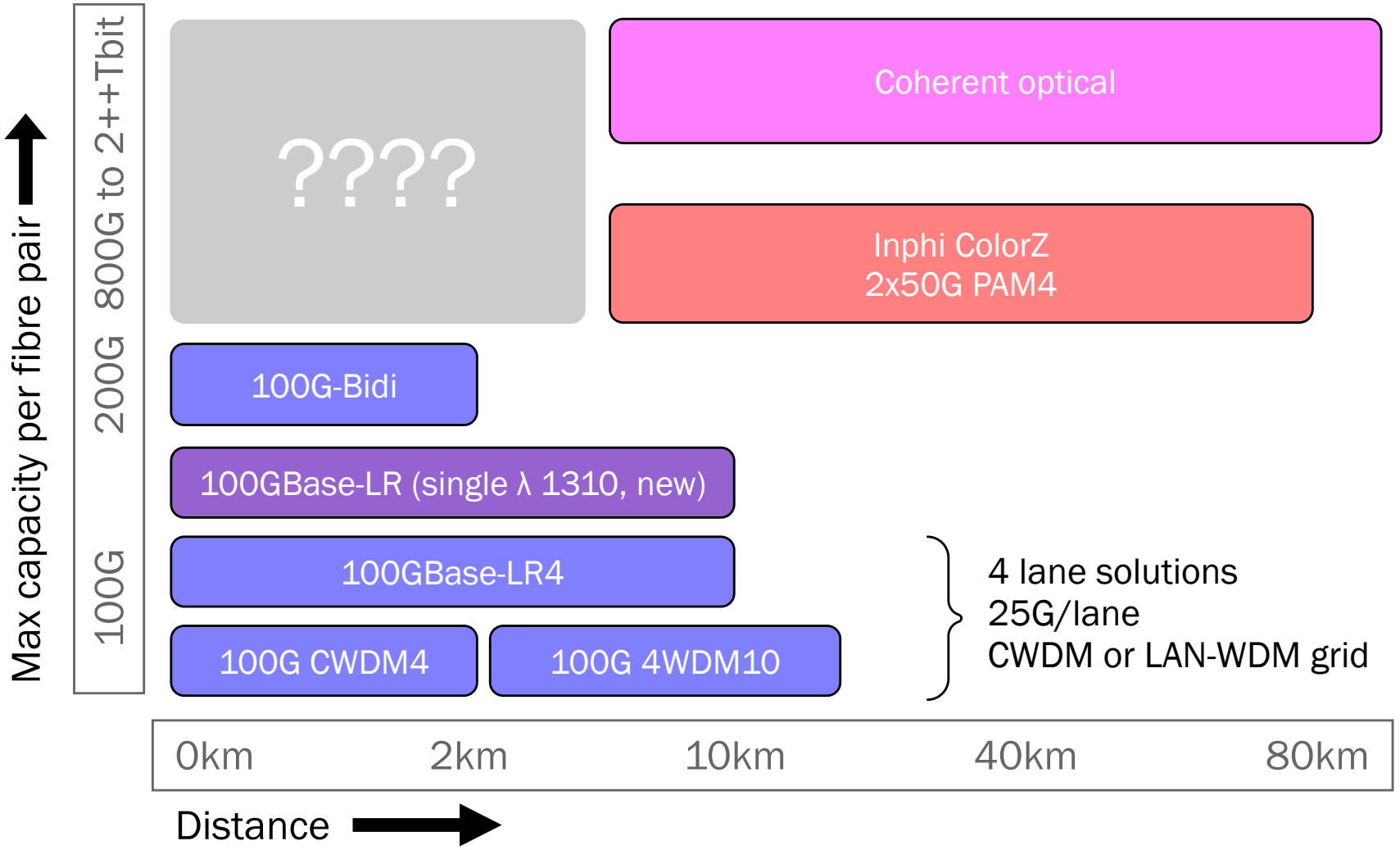
Open line systems

- Vendors start to produce dedicated systems to address these challenges
- Minimal configuration required
- “Feels like a passive DWDM mux” (but needs power)
- Contains optical amplification and dispersion compensation

Components in Smartoptics DCP-M series



100G era technologies



LONAP™

100G+ in the metro: Shorter links

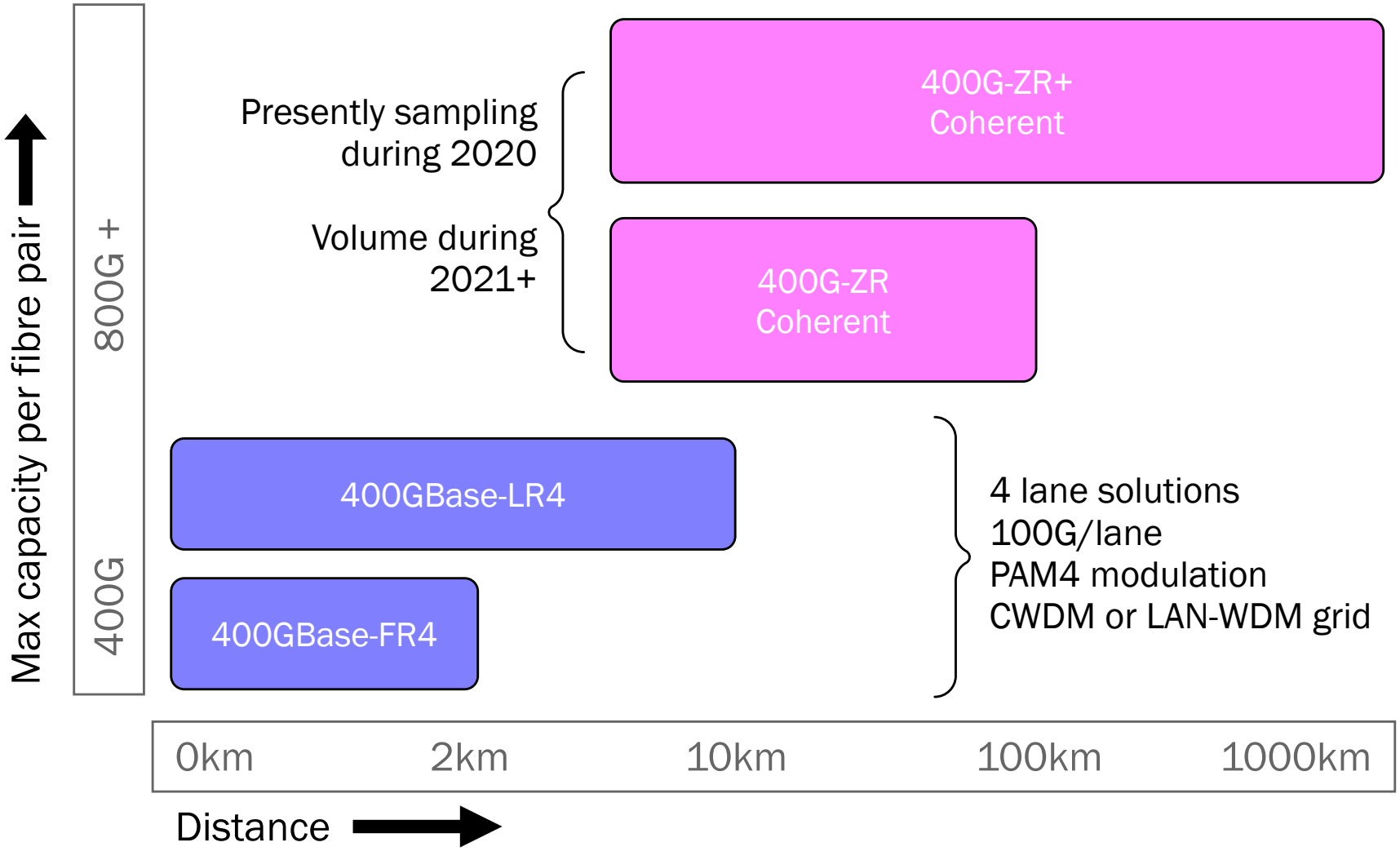
- Many of our links are in the < 10km range
- ”The grey area”
- Do cheaper options exist?
- Direct price comparison against MRC charges for additional dark fibre pairs and/or datacentre crossconnects

- Testing 100G-Bidi with APC connectors (Flexoptix)
  - functions OK for simple links < 2km
  - Scales to 200G/pair

400G

- 400G already shipping in quantity, for “LAN” links <10km
  - Growing very fast, bandwidth of deployed 400G ports is expected to exceed 100G this year
  - Optic cost is very competitive – per-gig cost already cheaper than 100G
- Obviously you will need 400G host ports
  - but that’s coming quite quickly for LONAP!
- In the future coherent ‘ZR’ and ‘ZR+’ pluggables will deliver distances from 10km up to thousands of km. Sampling during 2020.

400G era technologies



400G LAN options for SMF + 100G breakout

- 400GBase-FR4 (4x100 PAM4, CWDM grid, 2km reach)
- 400GBase-LR4 (4x100 PAM4, CWDM grid, 10km reach)

But also:

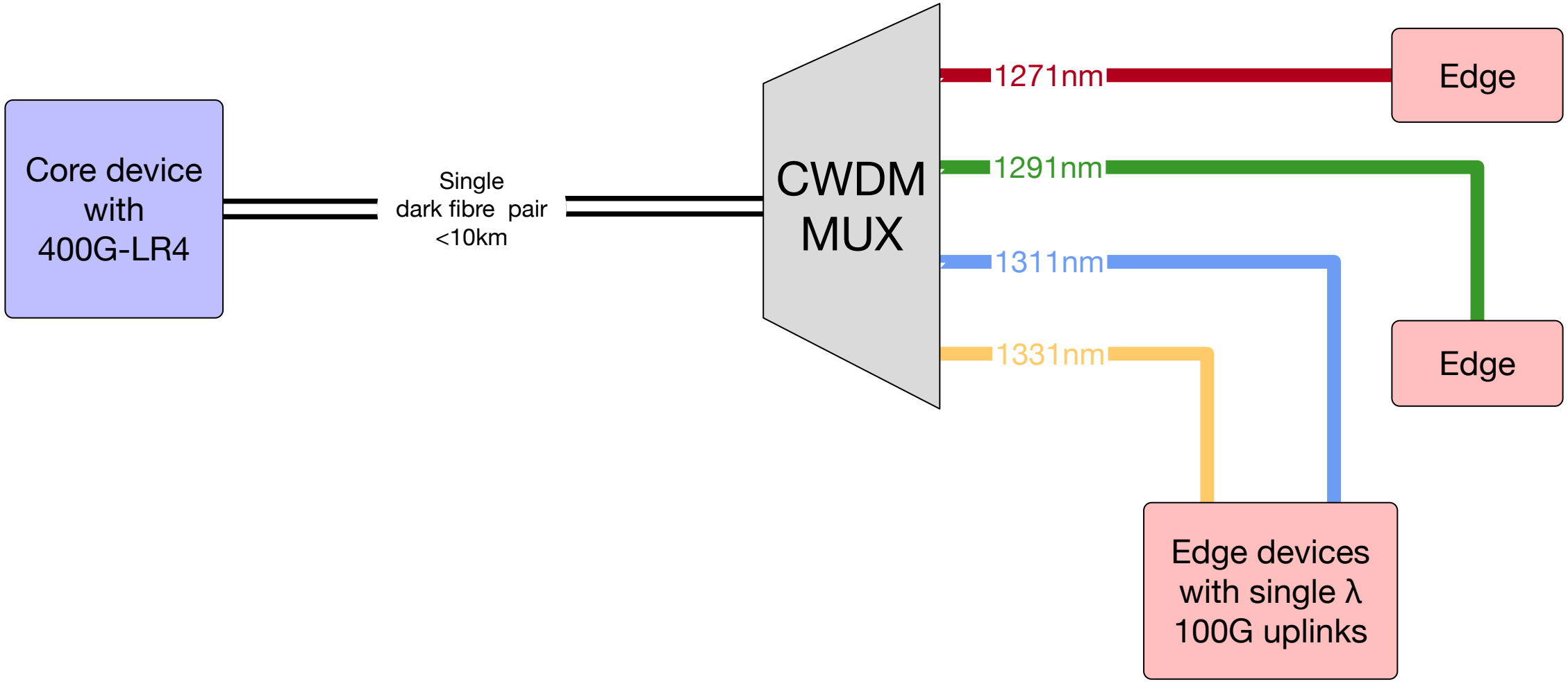
- 100GBase-LR (single λ 100G, 1311nm, 10km reach)
  - Interoperates with 400GBase-DR4 – breakout 4x100G lanes on an MPO connector
  - Already price-competitive vs. ”standard” 4-λ 100GBase-LR4

400G LAN options for SMF – CWDM breakout

- 400GBase-FR4 (4x100 PAM4, CWDM grid, 2km reach)
- 400GBase-LR4 (4x100 PAM4, CWDM grid, 10km reach)
- These optics use 4x100G lasers
- What if we could break these out into single 100G lanes?

  - Requires CWDM coloured 100G single-λ optics
  - Possible PoP architectures with 400G core and 100G-uplinked edge devices
  - Scaling existing 100G links to 4x100G
  - LONAP will test during 2H 2020

Possible CWDM 100G breakout architecture



Summary

- Optics world is fast-developing
- Some standards are moving quickly towards obsolescence, sometimes without even making it to manufacture
- Single-lambda 100G is likely to dominate shorter range interconnection in the future
- We can scale cost-effectively

Thoughts? Comments?