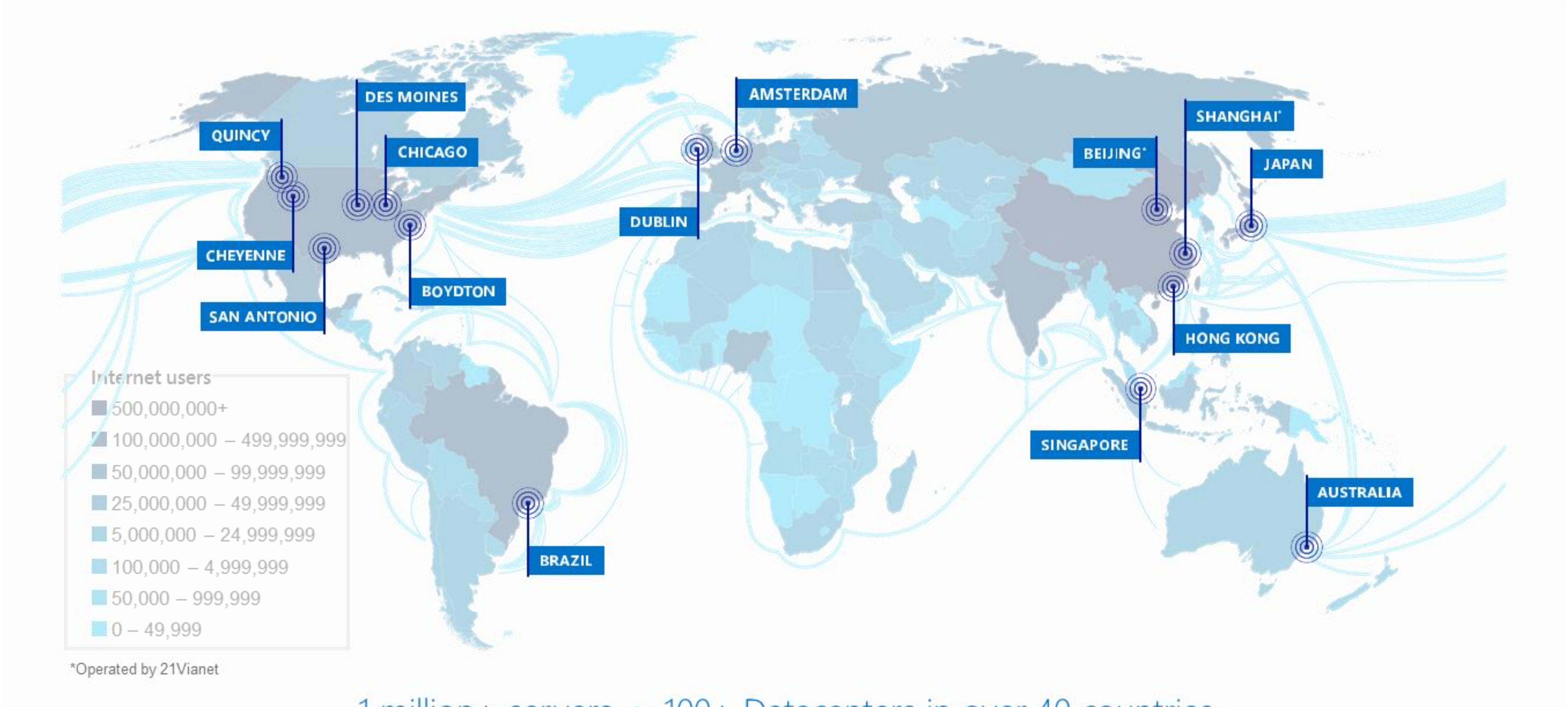
Metro Data Center Interconnect

Brad Booth
Azure Networking
July 2016

Microsoft

Microsoft's Global Data Center Footprint

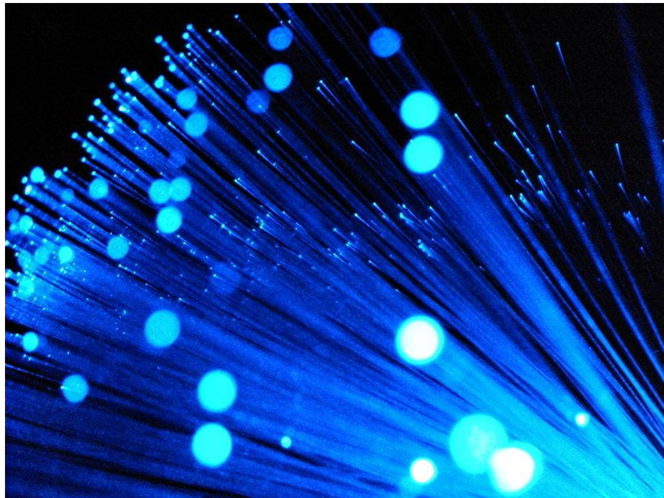


1 million+ servers • 100+ Datacenters in over 40 countries

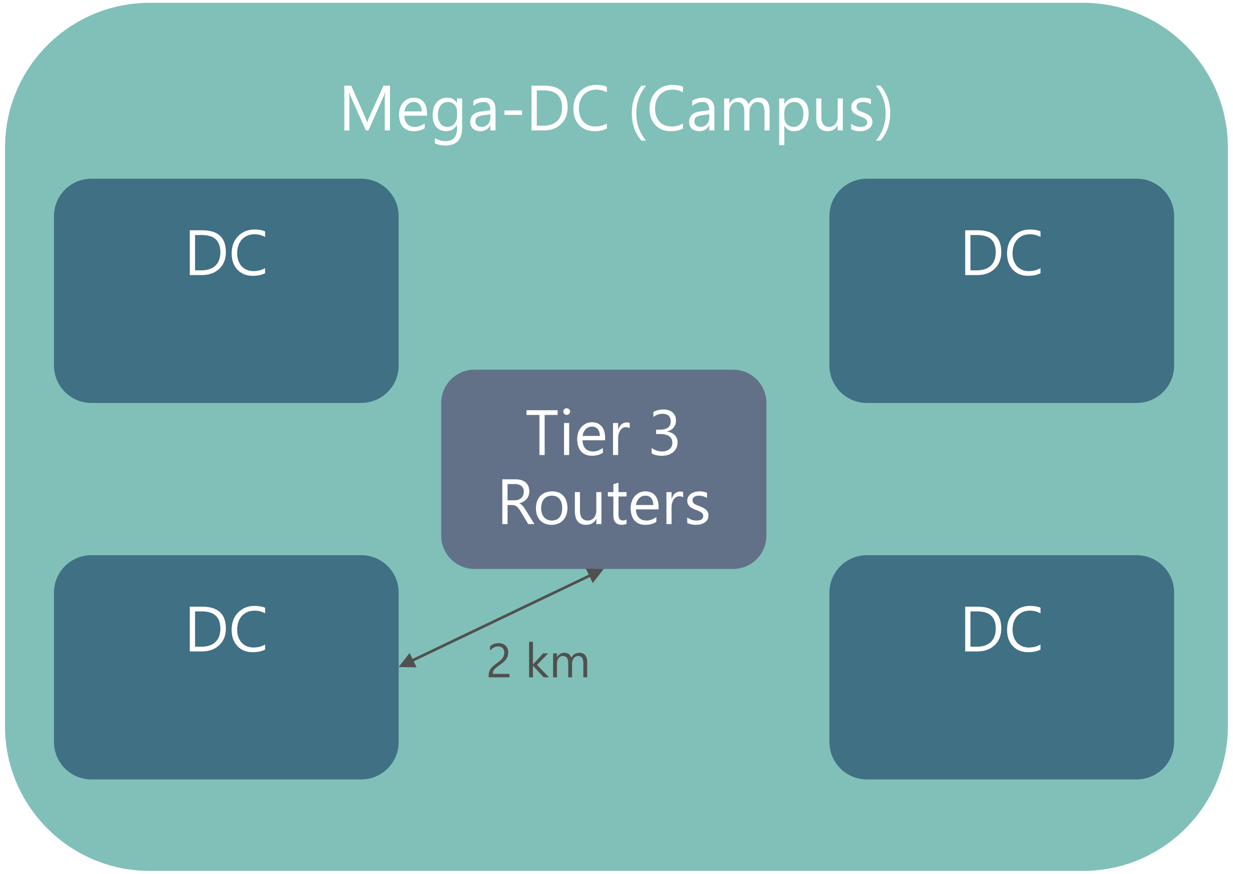
Microsoft’s Optical Network

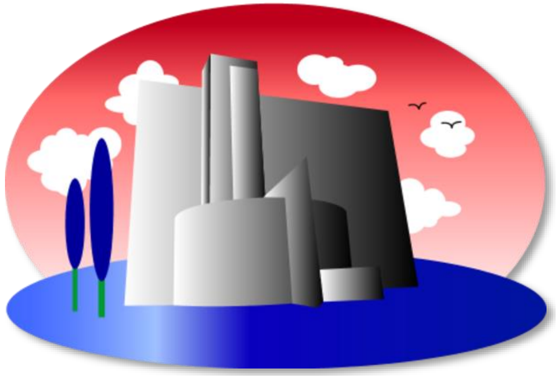
Microsoft owns and operates its own network infrastructure, including:

- **Intra-DC**: datacenters and all ‘inside datacenter’ hardware
- **Metro**: optical fiber and amplifiers / hardware connecting datacenters inside ‘regions’ (i.e. inside a city)
  - ≤ 70 km
- **Long-haul**: optical fiber and amplifiers / hardware connecting DCs between cities
  - 100 to 6,000 km
- **Subsea**: optical fiber cables
  - 5,000 to 10,000 km

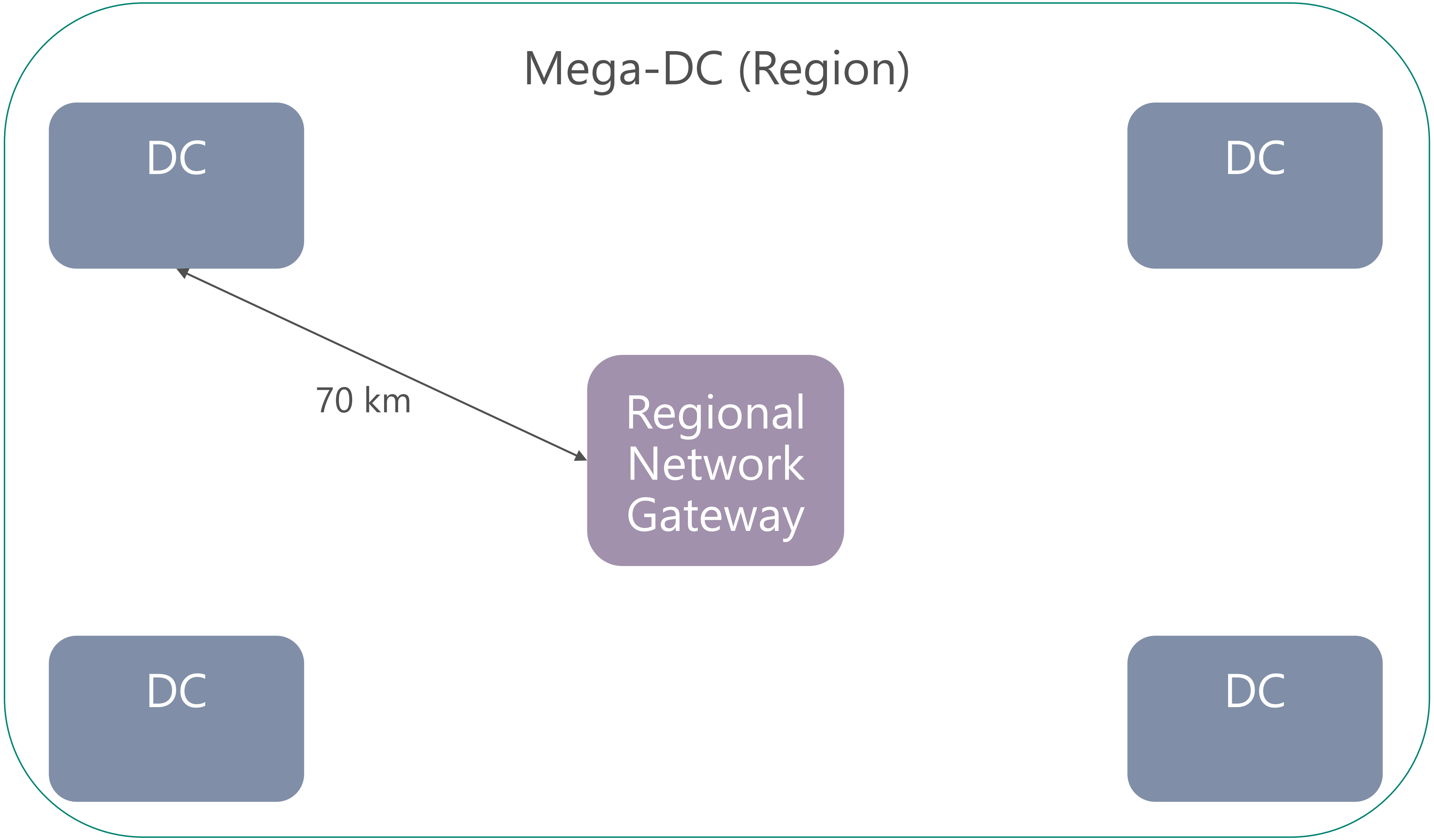


Mega Data Center Architecture v1.0





Mega Data Center Architecture v2.0



Campus → Region

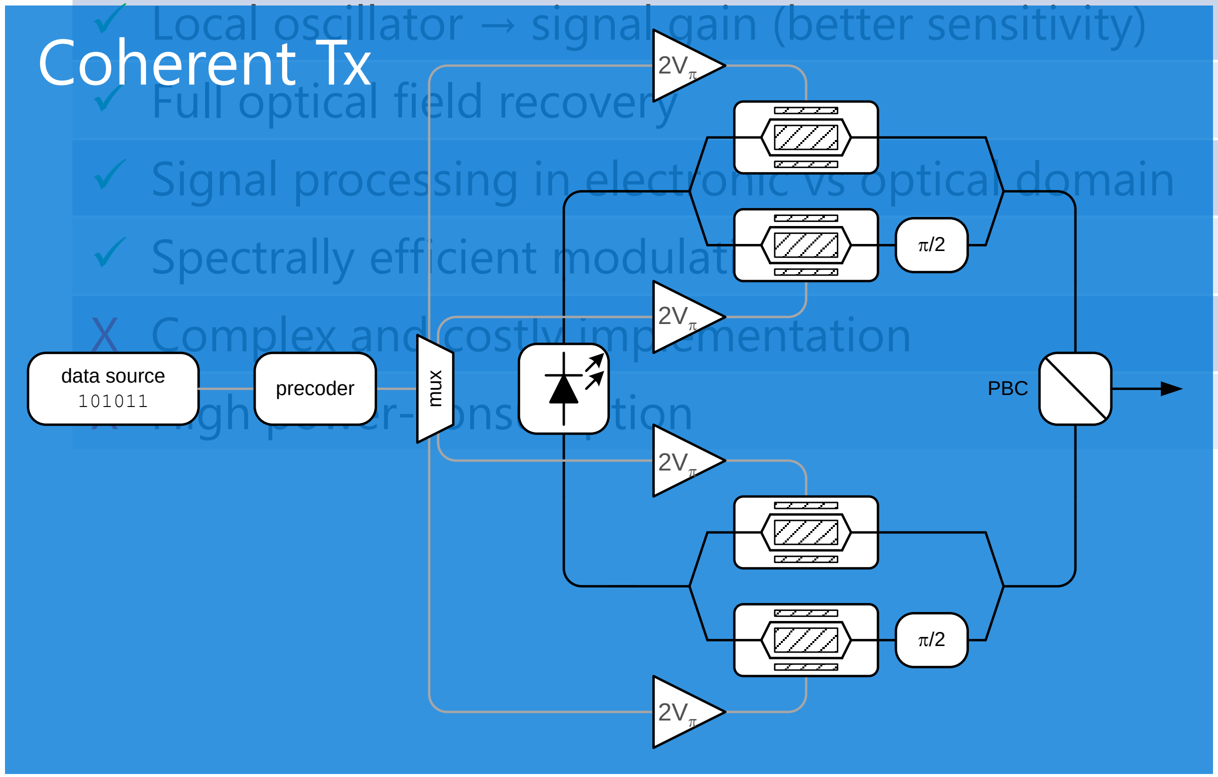
- Construction of a Mega-DC
  - V1.0 Campus doesn't meet pace of growth
  - V2.0 Region has reduced construction requirements; matches growth
- Reach requirements
  - 2 km maximum within a v1.0 Campus
  - 70 km maximum within a v2.0 Region
    - Latency restricted
- Fiber infrastructure
  - V1.0 Campus is fiber-rich parallel single-mode
  - V2.0 Region is fiber-limited single-mode

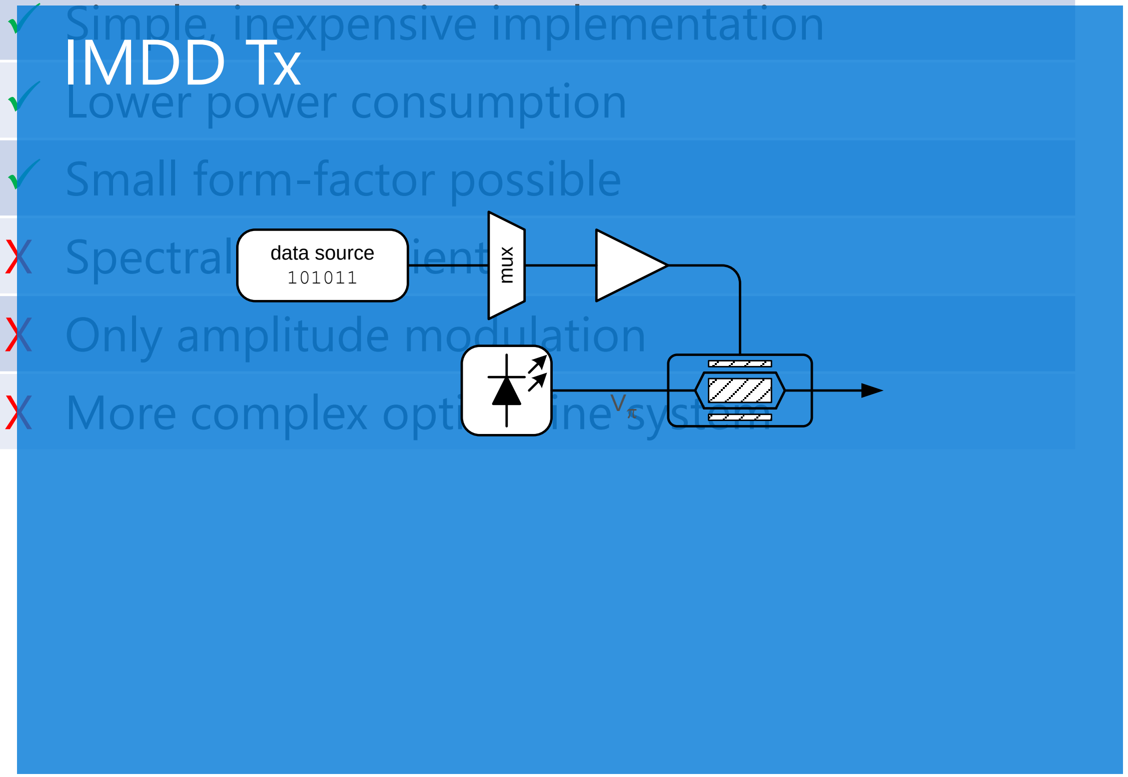
Modulation and Encoding

- Dimensions of encoding signals:
  - Intensity / Amplitude (e.g., ASK, or OOK)
  - Phase / Frequency (e.g., PSK)
  - Polarization – orientation of the light’s electric field
- Intensity-modulated direct detect (IMDD) vs Coherent
  - **IMDD**
    - Simplest implementation – only amplitude data retained because square-law detector
    - Limited performance and spectral efficiency
  - **Coherent**
    - Symbols transmitted with multiple bits per symbol
    - More complex implementation – mechanisms to act on amplitude, phase, and polarization
    - Highest performance and spectral efficiency

IMDD vs Coherent

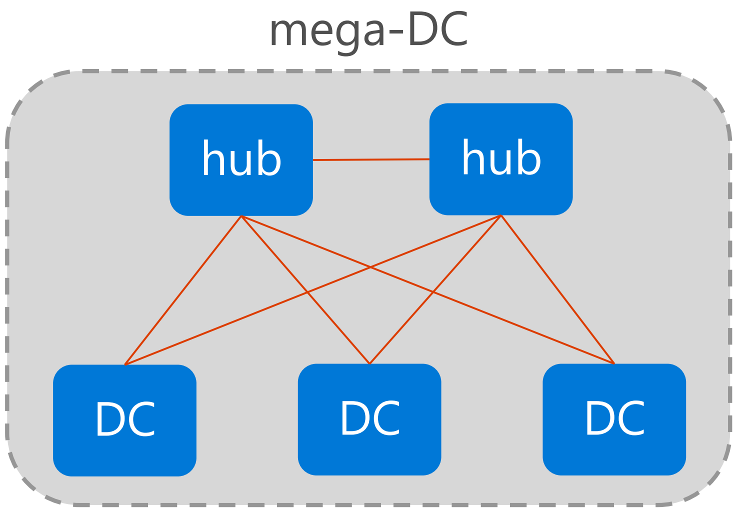
Coherent	IMDD
✓ Local oscillator → signal gain (better sensitivity)	✓ Simple, inexpensive implementation
✓ Full optical field recovery	✓ Lower power consumption
✓ Signal processing in electronic vs optical domain	✓ Small form-factor possible
✓ Spectrally efficient modulation	X Spectrally inefficient
X Complex and costly implementation	X Only amplitude modulation
X High power-consumption	X More complex optical line system



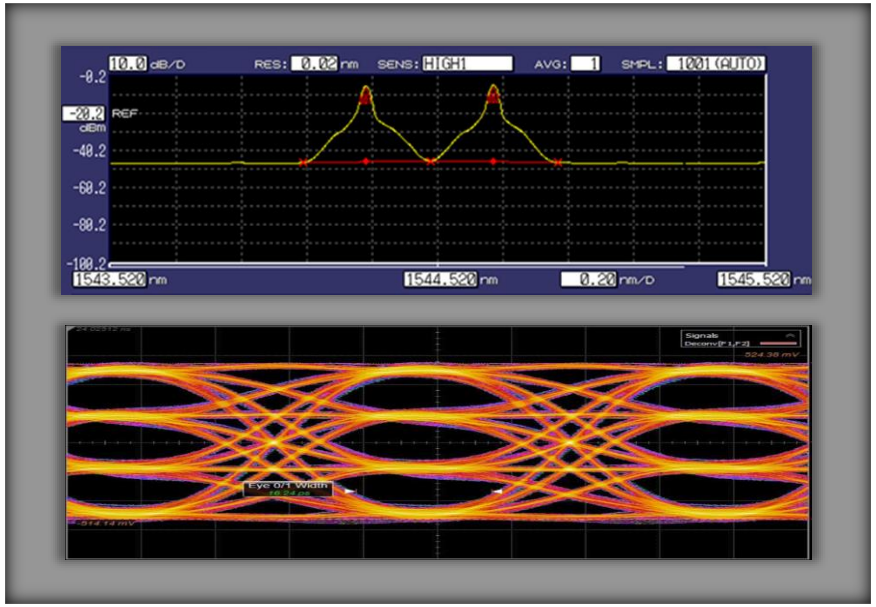


Madison Generation 1.0

- 100G per module
  - Dual-λ, 25 Gbaud PAM4 modulation
- Direct-detect demodulation (i.e. non-coherent)
- Silicon photonics-based optics
- FEC/DSP technology
- QSFP28 form-factor ≤ 4.5 W/plug



PAM4



Madison Advantages

- **Reduced space**
  - QSFP28 form-factor → 36 x 100G in a 1 RU line card
- **Reduced power consumption**
  - ~1/20th power of coherent
- **Reduced cost**
  - Lower cost than coherent (total cost of ownership over 3 years)
- **Elimination of Layer-1 interconnect**:
  - No optical interfaces between Ethernet switch and DWDM platform

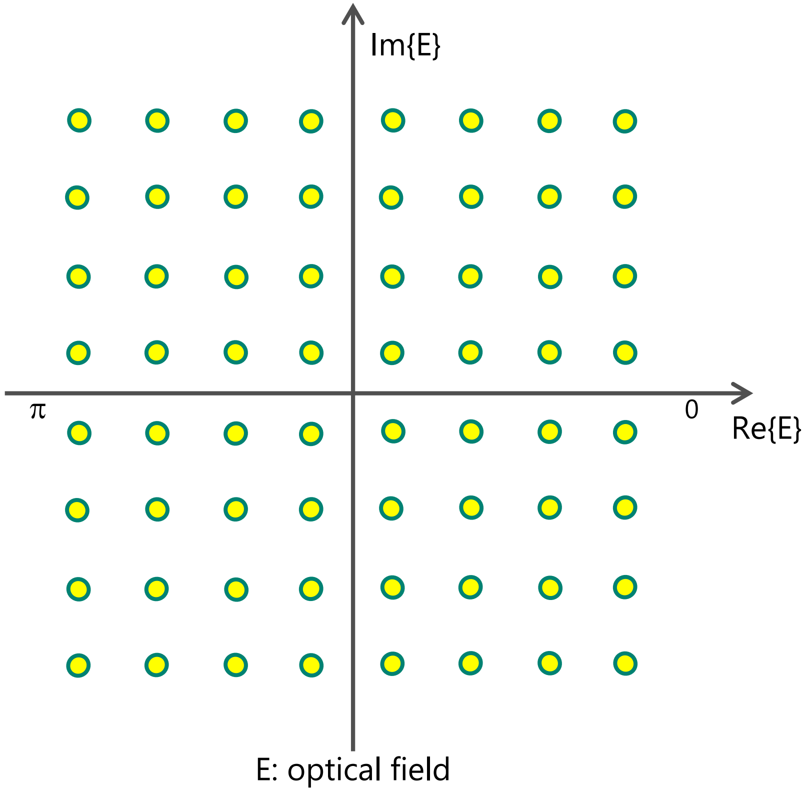
  → DWDM optics plug directly into the switch line card

Madison Generation 1.5

- Single-λ, 56 Gbaud PAM4 (100 Gb/s)
  - Capacity: 6.4 to 7.2 Tb/s (or maybe 9.6 Tb/s) per fiber
  - Cost target: parity with Madison 1.0
- Performance
  - Single laser (tunable?)
  - Better Tx and Rx power specs than Madison gen 1
  - More challenging OSNR and dispersion tolerance
  - Same line system components with minor changes to mux/demux scheme

Madison Generation 2

- "Coherent-lite" solution
- Current coherent chip sets
  - "One size fits all" metro and subsea is overkill for "DCI"
  - Space / cost / power is challenging to scale
- Solution
  - Eliminate subsea dispersion compensation
  - Focused baud rate and modulation format
  - Cost parity with Madison Gen 1
  - 38 Tb/s per fiber



<u>One possibility:</u>

400G 64QAM

~27dB rOSNR

Microsoft