CFI Consensus - Beyond 10km Optical PHYs

Consensus Presentation

John D'Ambrosia

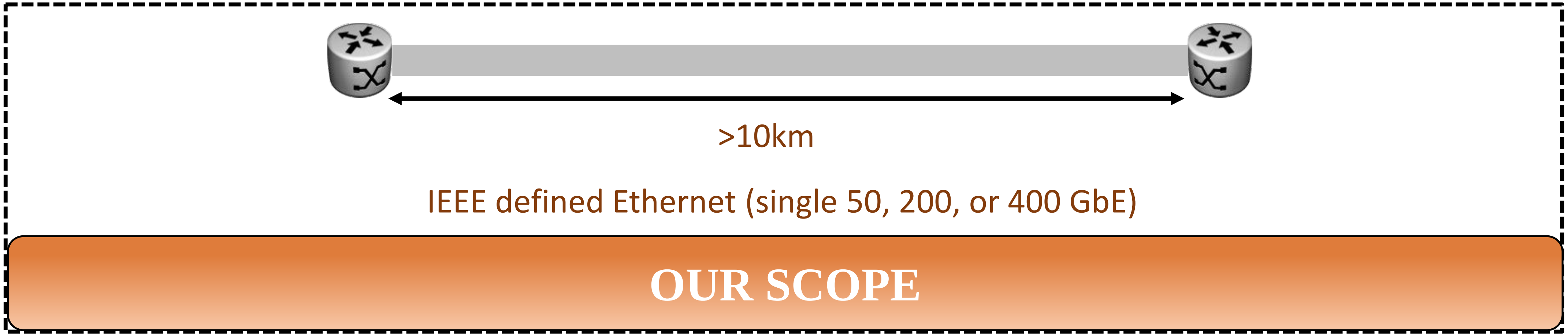
Futurewei, Subsidiary of Huawei

Objective for this Meeting

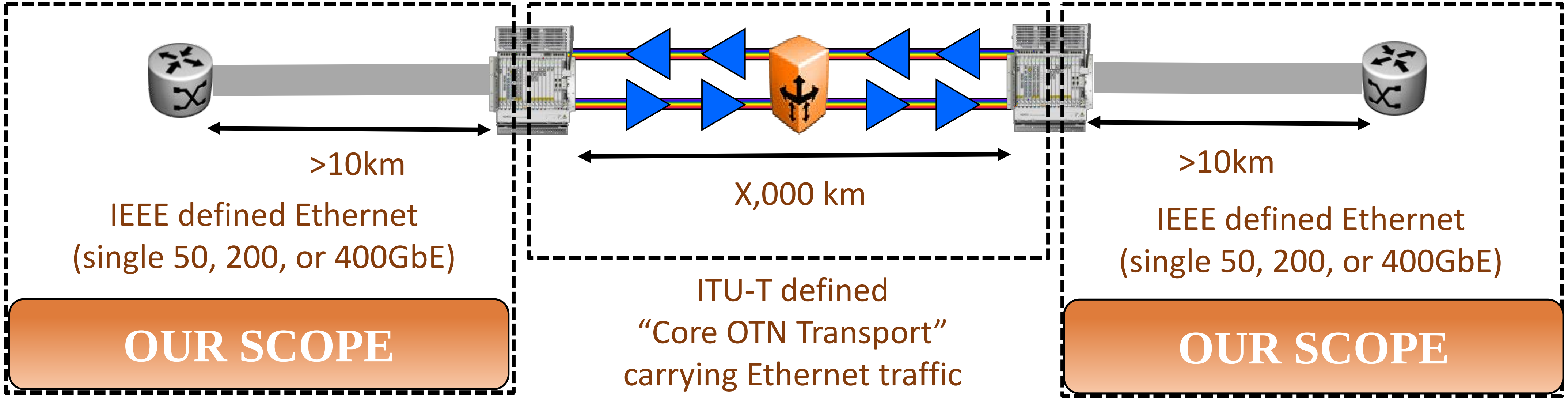
- **To *<u>measure the interest</u>* in starting a study group to address:**
  - **Beyond 10 km Optical PHYs for 50GbE, 200GbE, and 400GbE**
- **We don’t need to**
  - **Fully explore the problem**
  - **Debate strengths and weaknesses of solutions**
  - **Choose any one solution**
  - **Create PAR or five criteria**
  - **Create a standard or specification**
- **Anyone in the room may speak / vote**
- **RESPECT… give it, get it**

What Are We Talking About?

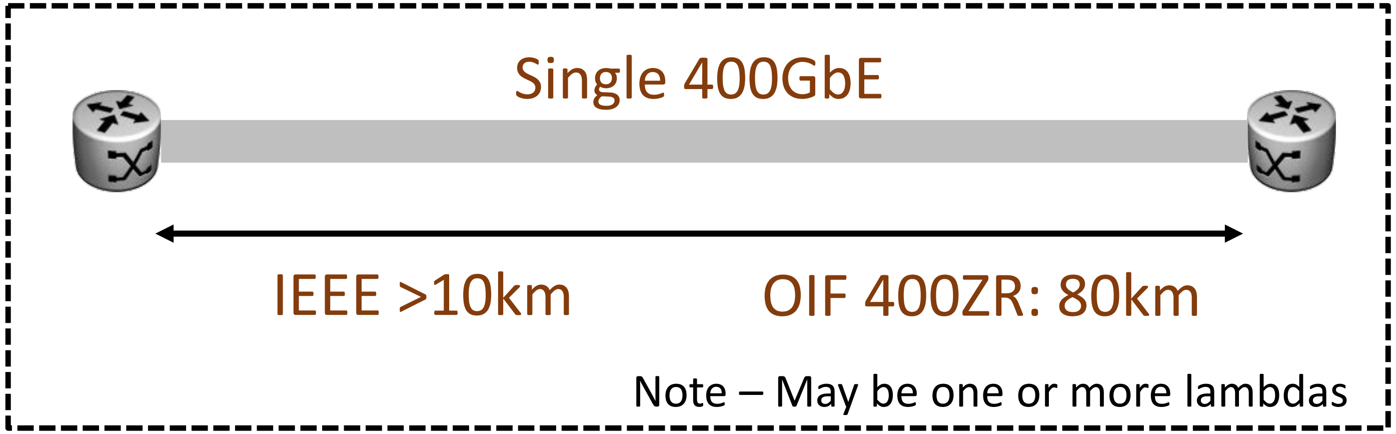
Scenario #1

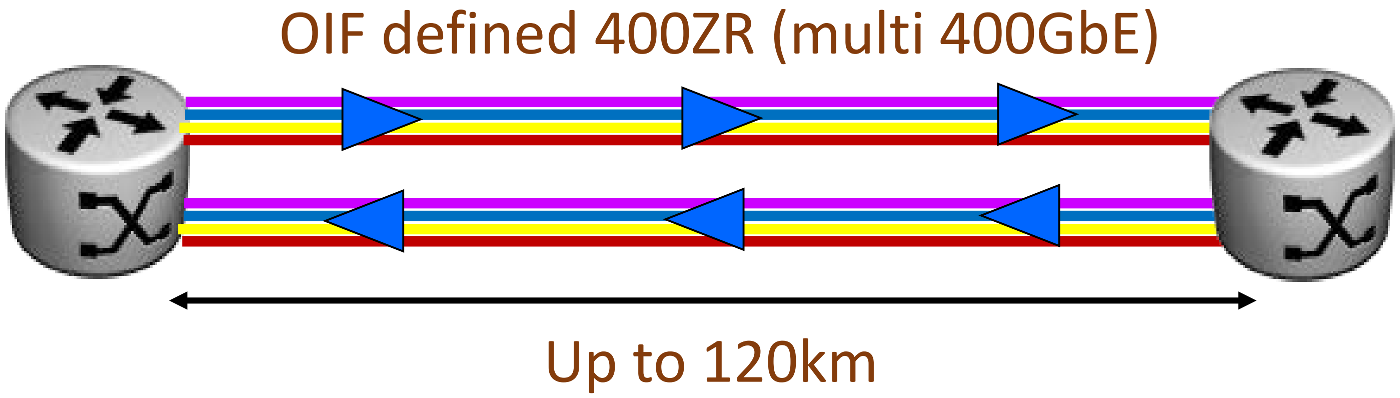


Scenario #2



400GbE and Potential Relationship to OIF 400ZR Data Center Interconnect (DCI) Solution





- **Coherent Optics is one potential solution to achieving reaches beyond 10km for 400GbE.**
- **It is not within the proposed scope of this effort to do a multi 400GbE coherent optical solution.**
- **It is recognized that a coherent solution developed by either organization could be leveraged for both application spaces.**

Agenda

- **Addressing Reaches Beyond 10km**
  - **John D'Ambrosia, Futurewei, Subsidiary of Huawei**
- **The Technical Aspect- Beyond 10km Optical PHYs**
  - **David Lewis, Lumentum**
  - **Tom Williams, Acacia**
- **Why Now?**
  - **John D'Ambrosia, Futurewei, Subsidiary of Huawei**
- **Straw Polls**

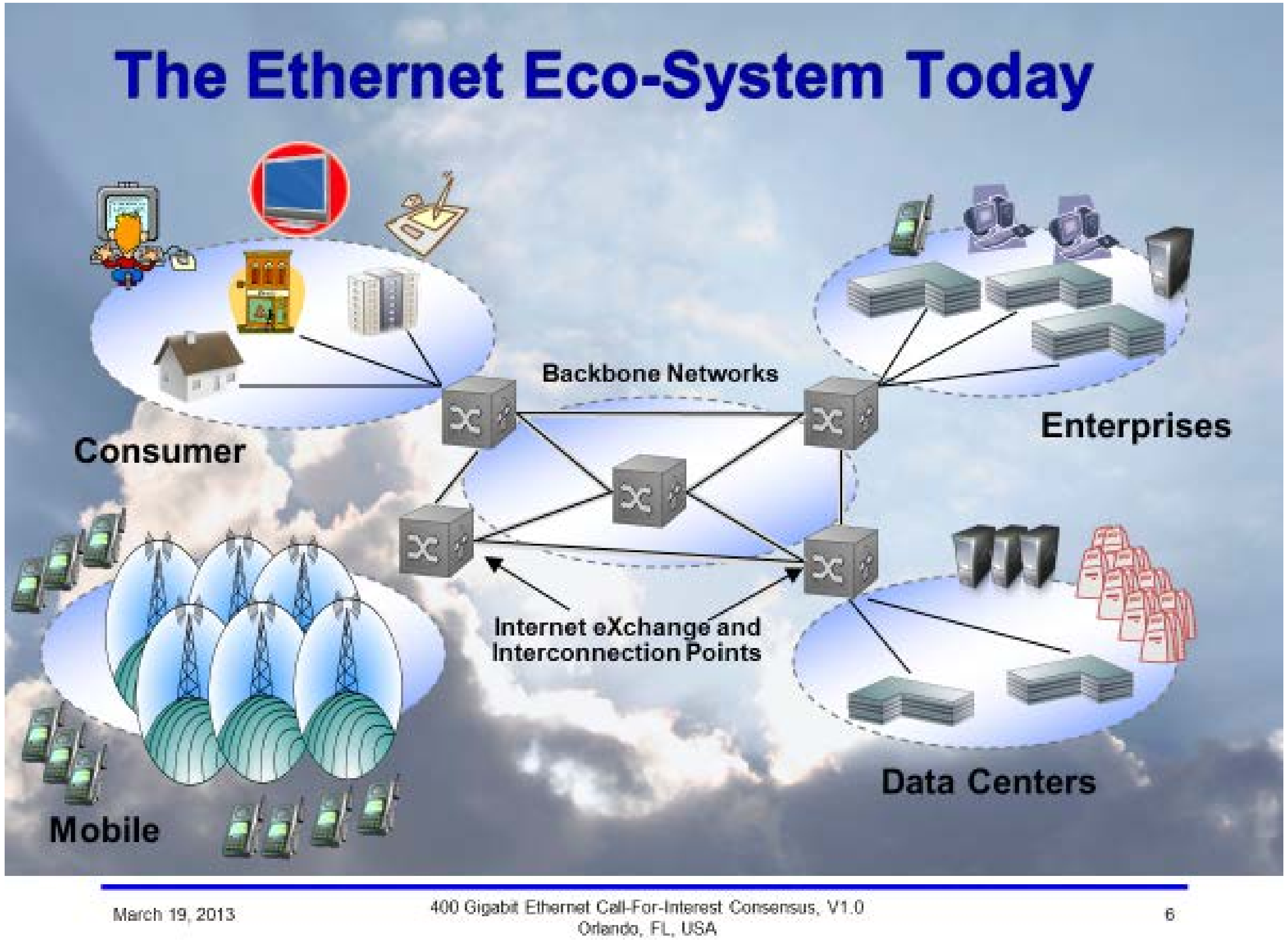
Addressing Reaches Beyond 10km

Today's Point-to-Point SMF Ethernet Family

	500m	2km	10km	20km	40km
10GBASE-			L4		ER
25GBASE-			LR		ER
40GBASE-	PSM4	FR	LR4		ER4
50GBASE-		FR	LR		
100GBASE-	PSM4 DR	10X10 CWDM4 / CLR4	LR4 / WDM4-10	WDM4-20	ER4 / WDM4-40
200GBASE-		FR4	LR4		
400GBASE-	DR4	FR8	LR8		

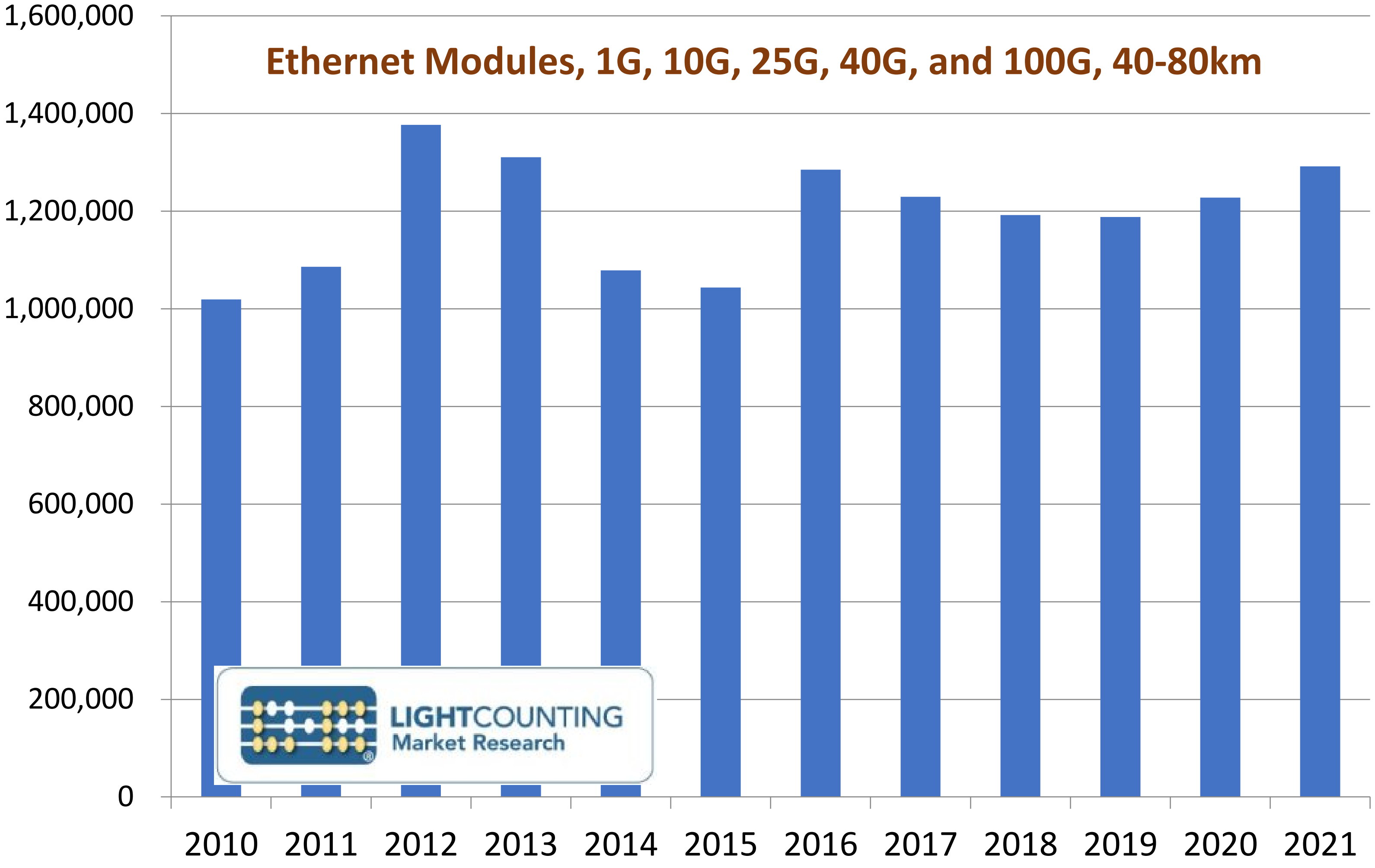
Black Text	IEEE Standard
Red Text	In Standardization
Blue Text	Non-IEEE standard but complies to IEEE electrical interfaces

Beyond 10km Optics Throughout The Eco-System



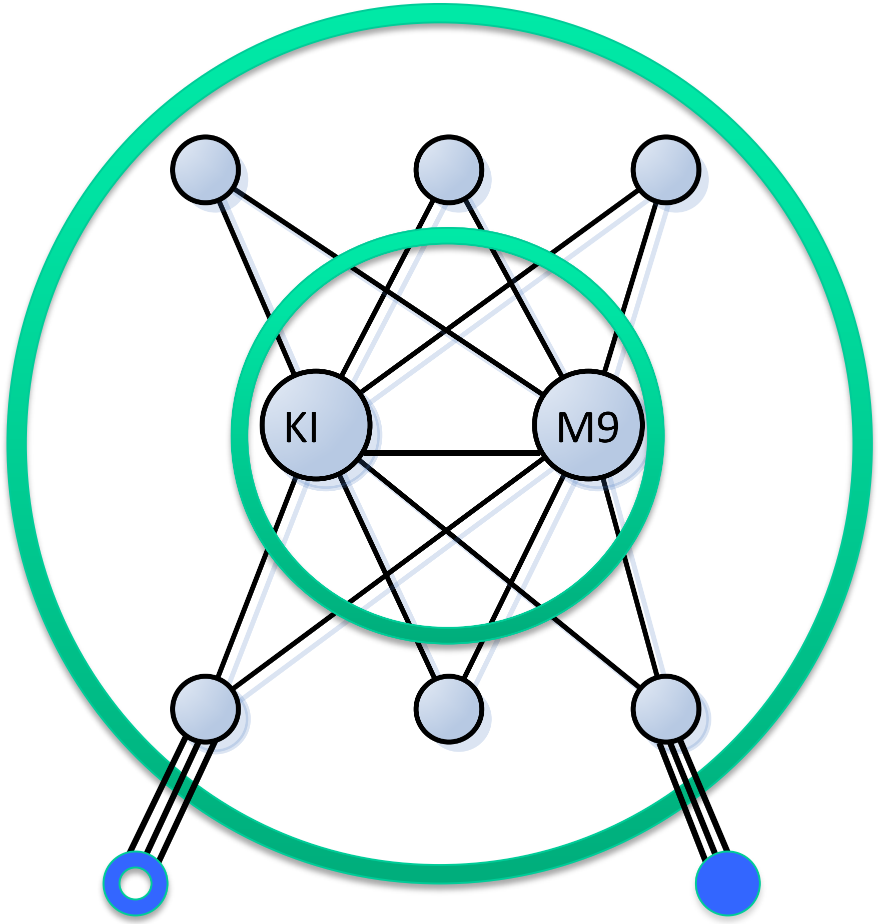
- **Not “Data Center”**
- **Exists throughout the Eco-System**
- **3 Million units for 40km and beyond shipped annually (see next page)**
- **Continuing bandwidth growth factors resonate throughout the ecosystem**
- **Not targeted by Ethernet standards for 50GbE, 200GbE, and 400GbE**

Annual Shipments for 40km+ Applications



- For 100GbE, 40km, LightCounting projects a market that will roughly triple in value from 2017 to 2021.
- SONET 40-80km shipments represent another half-million units in 2016. SONET is transitioning to Ethernet.
- 1 / 2.5 / 10 Gb/s DWDM / CWDM 40km & 80km optics will exceed 1M units this year and growing
- Totals are for merchant supplier shipments. Captive supply could add another half-million units.
- Data courtesy of LightCounting

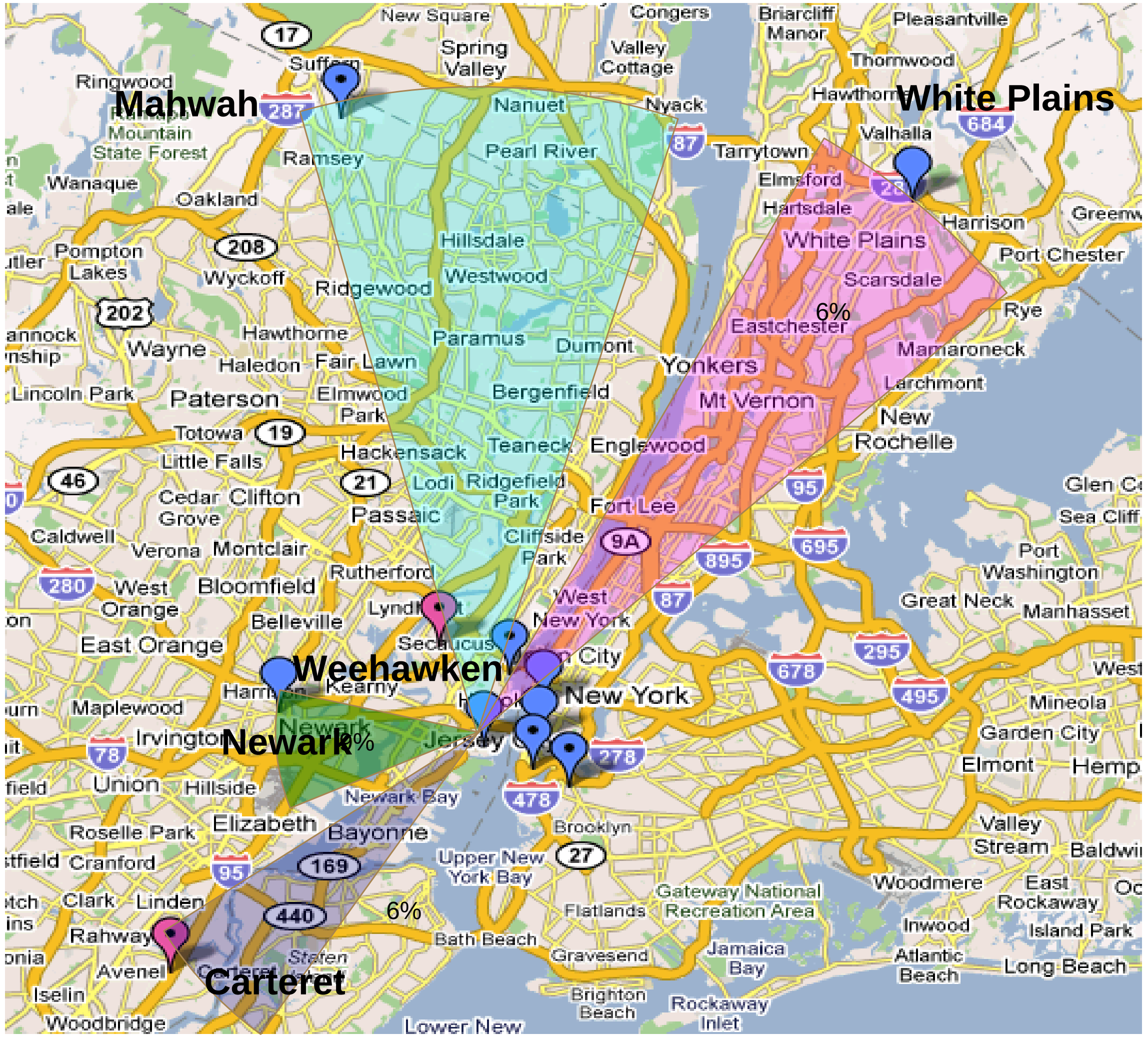
MSK-IX & Geographical Challenges



Source: Alexander Ilin, MSK-IX

- MLAG interaction between KI & M9 (~40km distance)
- Passive 10G DWDM solution between core, predictable network size
- Smooth migration from old equipment to a new one
- Ring-topology concept:
  - Tier 0 – connect core to each other,
  - Tier 1 – core datacenters and switches,
  - Tier 2 – edge datacenters.
- Current capacity between several Tier1 switches and Core: 640Gbps (n x 10G) with Future plans 100G+ links between them.
- Need solution for 100G+ optical transceivers between Core & Tier1 up to 40 km

NY, USA Financial Industry & Geographical Challenges



Note: All locations a for illustration purposes only and do not reflect actual locations

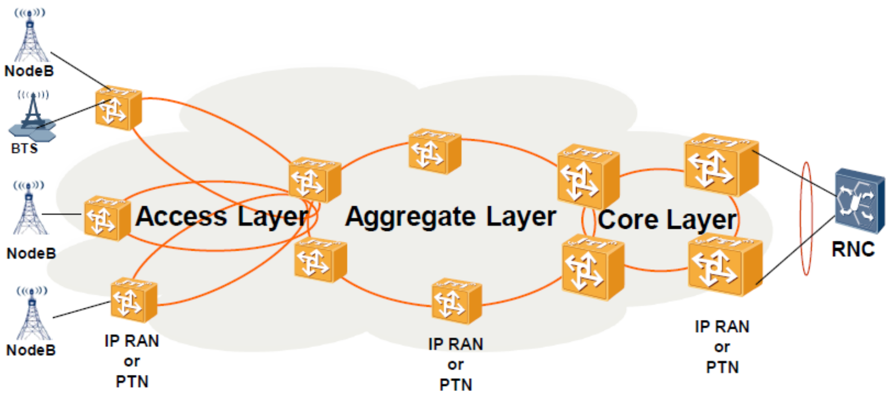
Weehawken, NJ

- **Carrier Access (Global) / Co-location Facility**
- **Used extensively by Financial Industry to support:**
  - **Connections to carrier access & hosting centers**
  - **Connections to "Execution Venues"**
- **Connections (Line of Sight)**
  - **Newark, NJ (16 km)**
  - **Carteret, NJ (32 km)**
  - **White Plains, NY (40 km)**
  - **Mahwah, NJ (42km)**

Source: Andrew Bach, Independent

Mobile Backhaul Demand for Beyond 10km

40km Reach in Mobile Backhaul Network



- In huang_ecdc_01_0716 and observation from shipment in Carrier network, 40km volume is increasing

Statistics for 10GE & 100GE Modules used in PTN, as of June, 2016				
Transmission Distance	<2km	10km	40km	80km
10GE distribution	0.28%	44.46%	44.05%	11.20%
100GE distribution (more than 15K modules)	0	56.43%	34.59%	8.97%

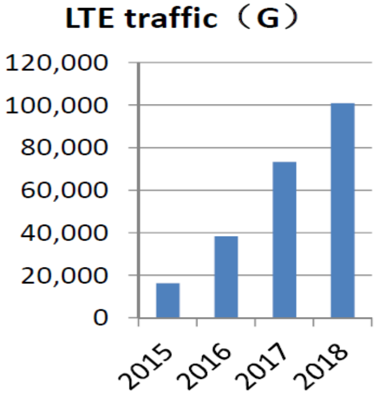
HUAWEI TECHNOLOGIES CO., LTD. Page 7 HUAWEI

Present status and forecast

- According to our survey, long distance module is a mandatory requirement for us

Statistics for 10GE & 100GE Modules used in PTN, as of June, 2016				
Transmission Distance	<2km	10km	40km	80km
10GE distribution	0.28%	44.46%	44.05%	11.20%
100GE distribution (more than 15K modules)	0	56.43%	34.59%	8.97%

- According to the increase of LTE traffic, as LTE backhaul network, PTN will face 4~5 times traffic in 2017 or 2018.
- Then we will have to use 400GE interface in the same scenario and take the same percentage with 100GE and 10GE.
- In 2018~2019, we expected the requirement for 400GE ER modules will be more than 10K.

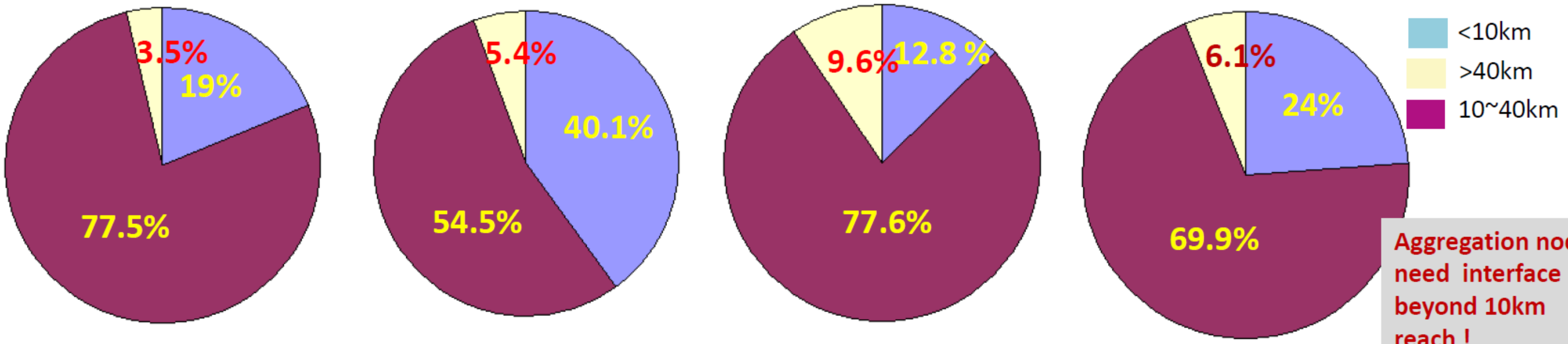


Source: Xinyuan Wang Huawei, http://www.ieee802.org/3/ad_hoc/ngrates/public/16_09/wang_ecdc_01_0916.pdf

Source: Huang/ Cheng, China Mobile, http://www.ieee802.org/3/ad_hoc/ngrates/public/16_07/huang_ecdc_01_0716.pdf

Aggregation node distance from actual networks

As metro core usually use WDM/OTN to extend reach distance of Ethernet interface, therefore current aggregation layer transmission distance is crucial to the future higher bitrate interface, such as 200GE and 400GE,etc.
Furthermore, each metro network may has its own distribution characteristic of reach distance, and some metro aggregation layer node distance from actual networks in China are investigated, and these nodes would has the requirement to deploy link capability more than 10GE.



Province A some metro aggregation layer reach distribution

Province B some metro aggregation layer reach distribution

Province C some metro aggregation layer reach distribution

Some metro aggregation layer reach distribution (ave.)

Aggregation node need interface beyond 10km reach !

IEEE 802.3 NG-ECDC Ad Hoc, July, 2016 , San Diego

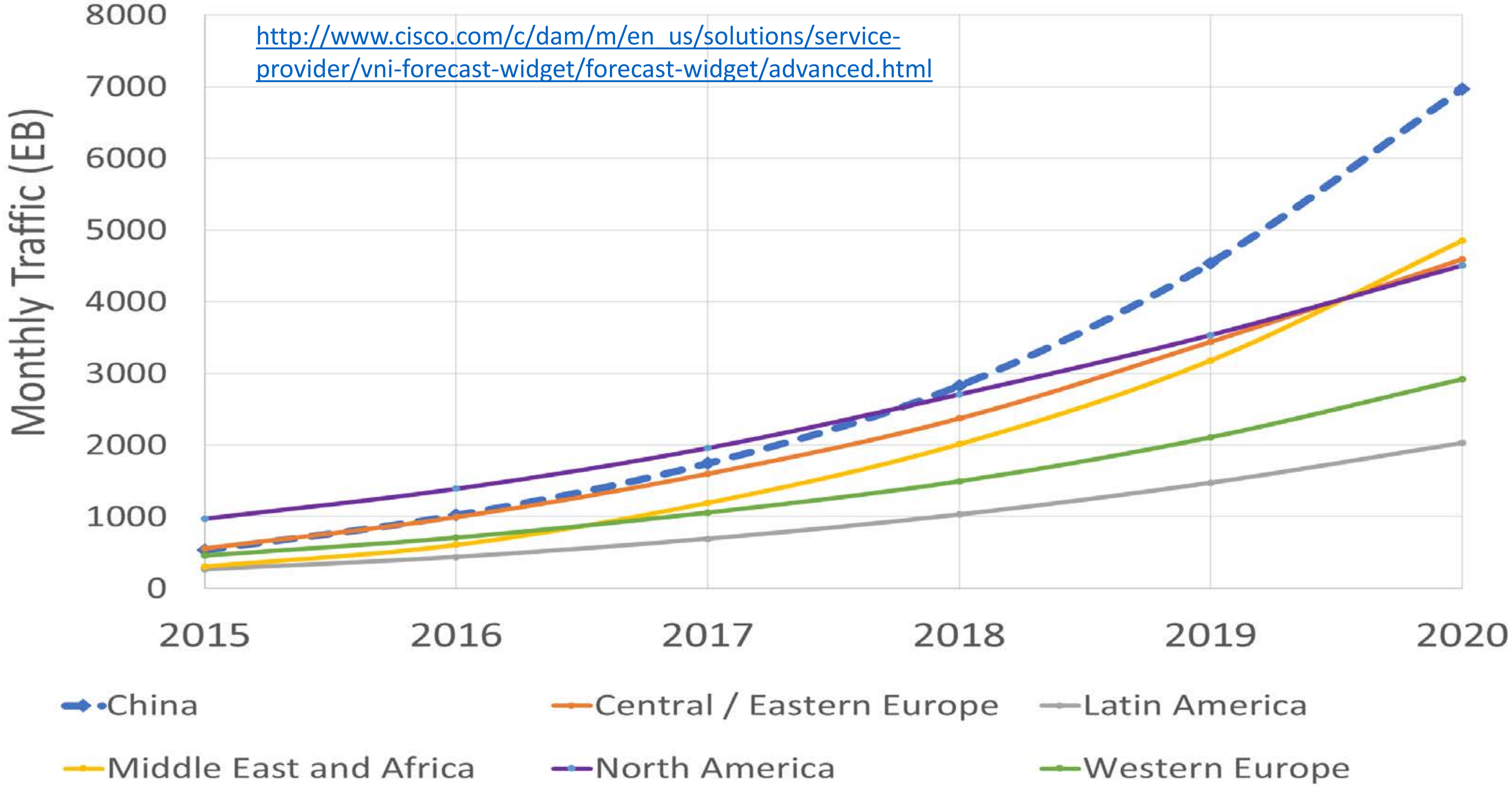
Source: Wenyu Zhao, CAICT
http://www.ieee802.org/3/ad_hoc/ngrates/public/16_07/zhao_ecdc_01_0716.pdf

Summary Observed Reaches - Telecom

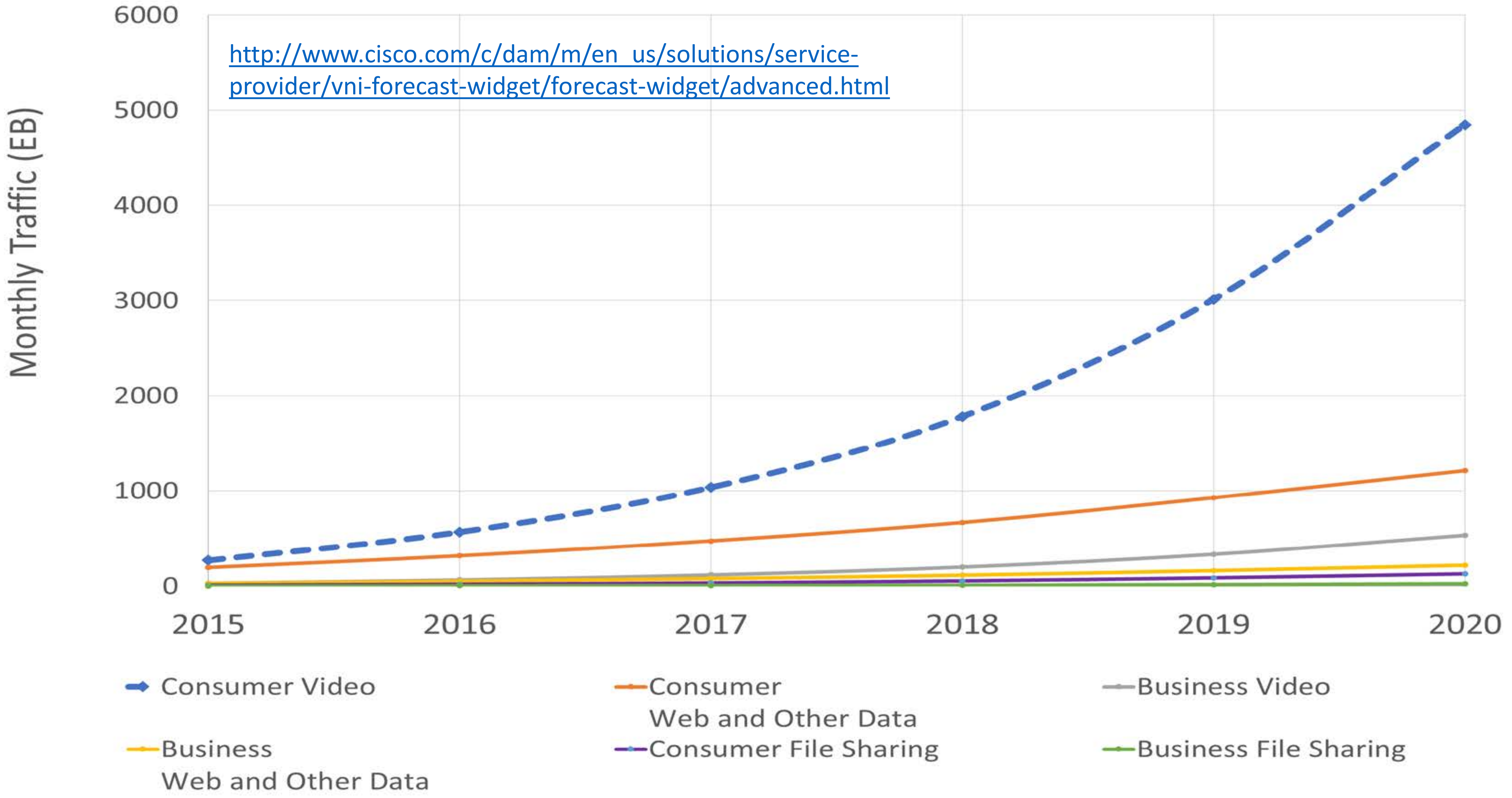
Source		<2km	10km	40km	>40km	80km
China Mobile *	10GbE	0.3%	44.5%	44.1%	-	11.2%
	100GbE	0	56.4%	34.6%	-	9.0%
CAICT Aggregation Nodes ** (200GbE / 400GbE)	Province A	-	19.0%	77.5%	3.5%	-
	Province B	-	40.1%	54.5%	5.4%	-
	Province C	-	12.8%	77.6%	12.8%	-
	Province D	-	24%	69.9%	6.1%	-
LightCounting	10 GbE All	- ***	93%	5.4%	-	1.6%
	10 GbE Telecom	0	76%	17%	-	7%

* - Source: Huang/ Cheng, China Mobile, http://www.ieee802.org/3/ad_hoc/ngrates/public/16_07/huang_ecdc_01_0716.pdf
** - Source: Wenyu Zhao, CAICT< http://www.ieee802.org/3/ad_hoc/ngrates/public/16_07/zhao_ecdc_01_0716.pdf
*** - 10GLR “Subspec” volume not included for this analysis

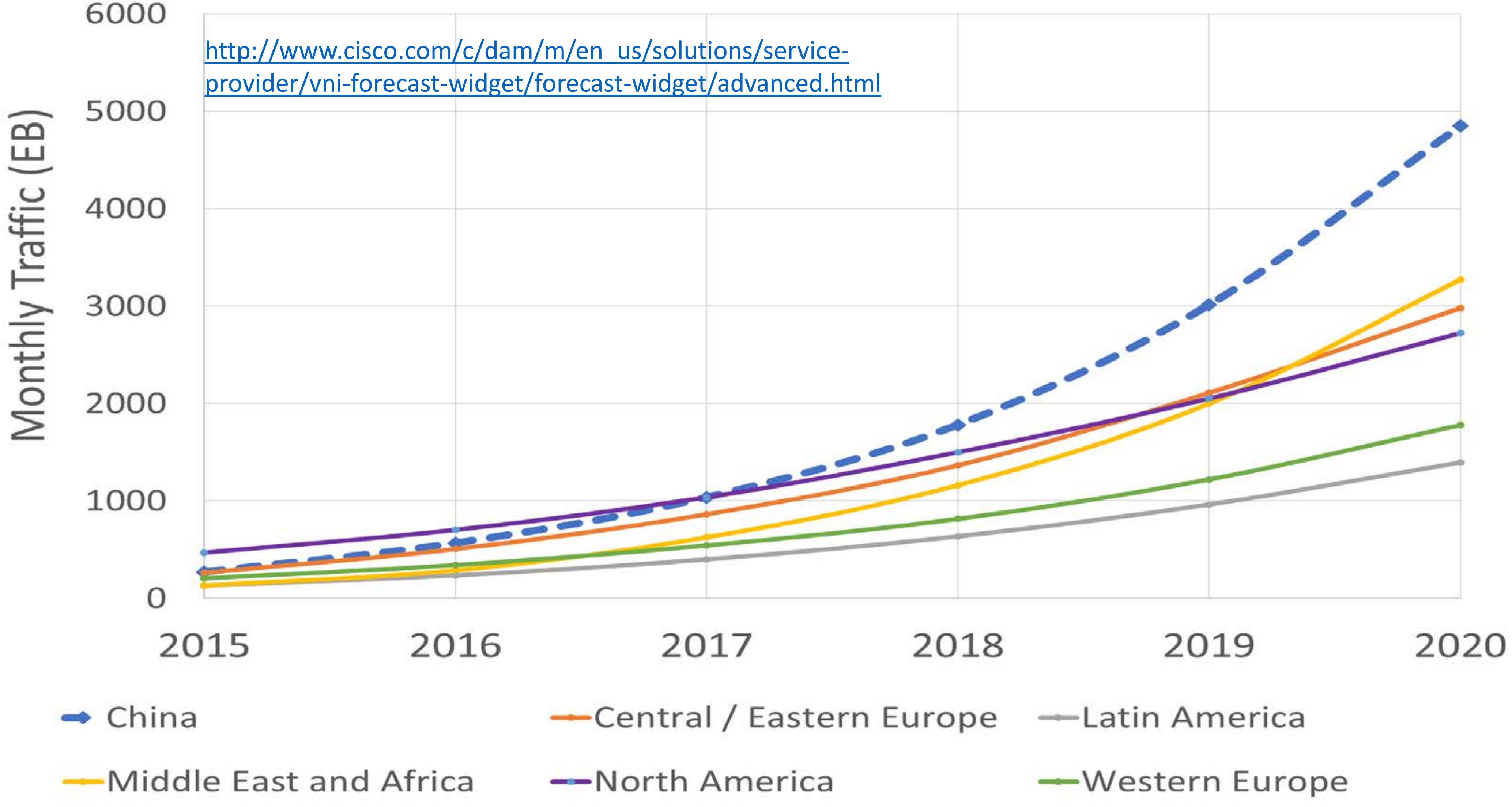
Mobile Networks Bandwidth Trends



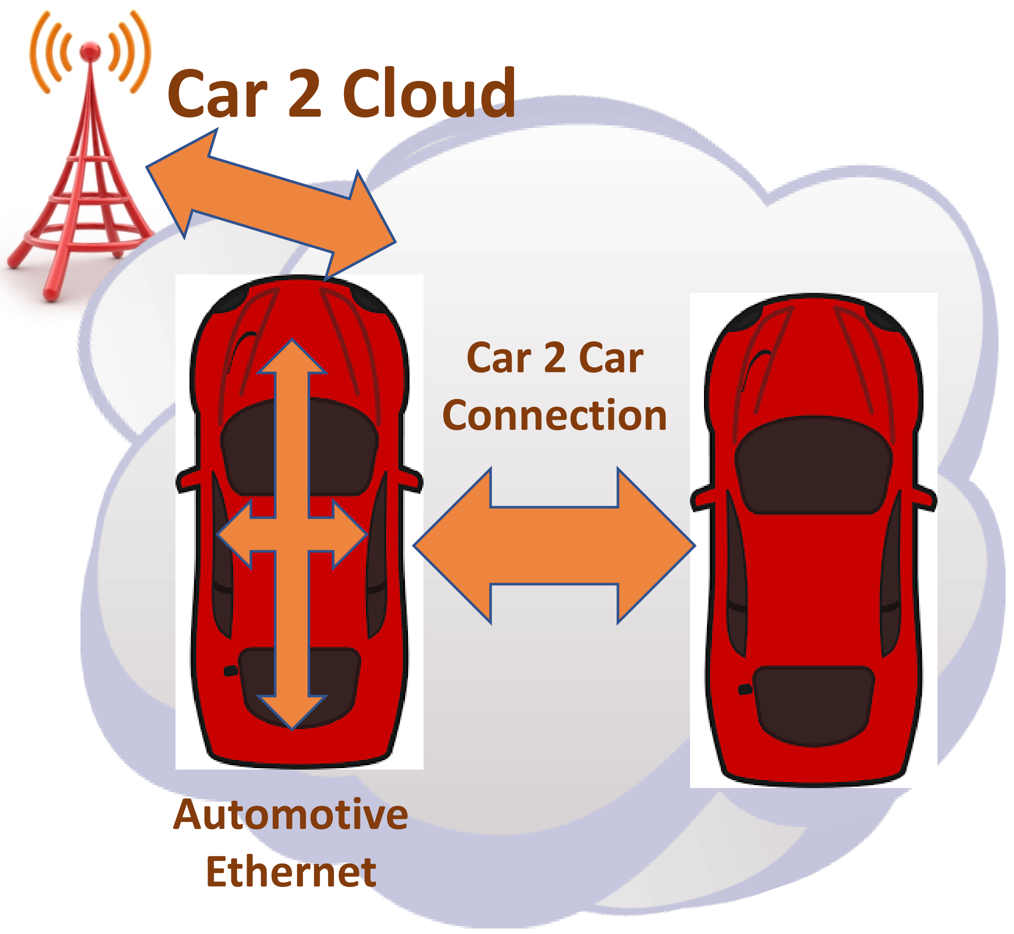
Mobile Networks - Application Bandwidth - China



Mobile Networks - Consumer Video

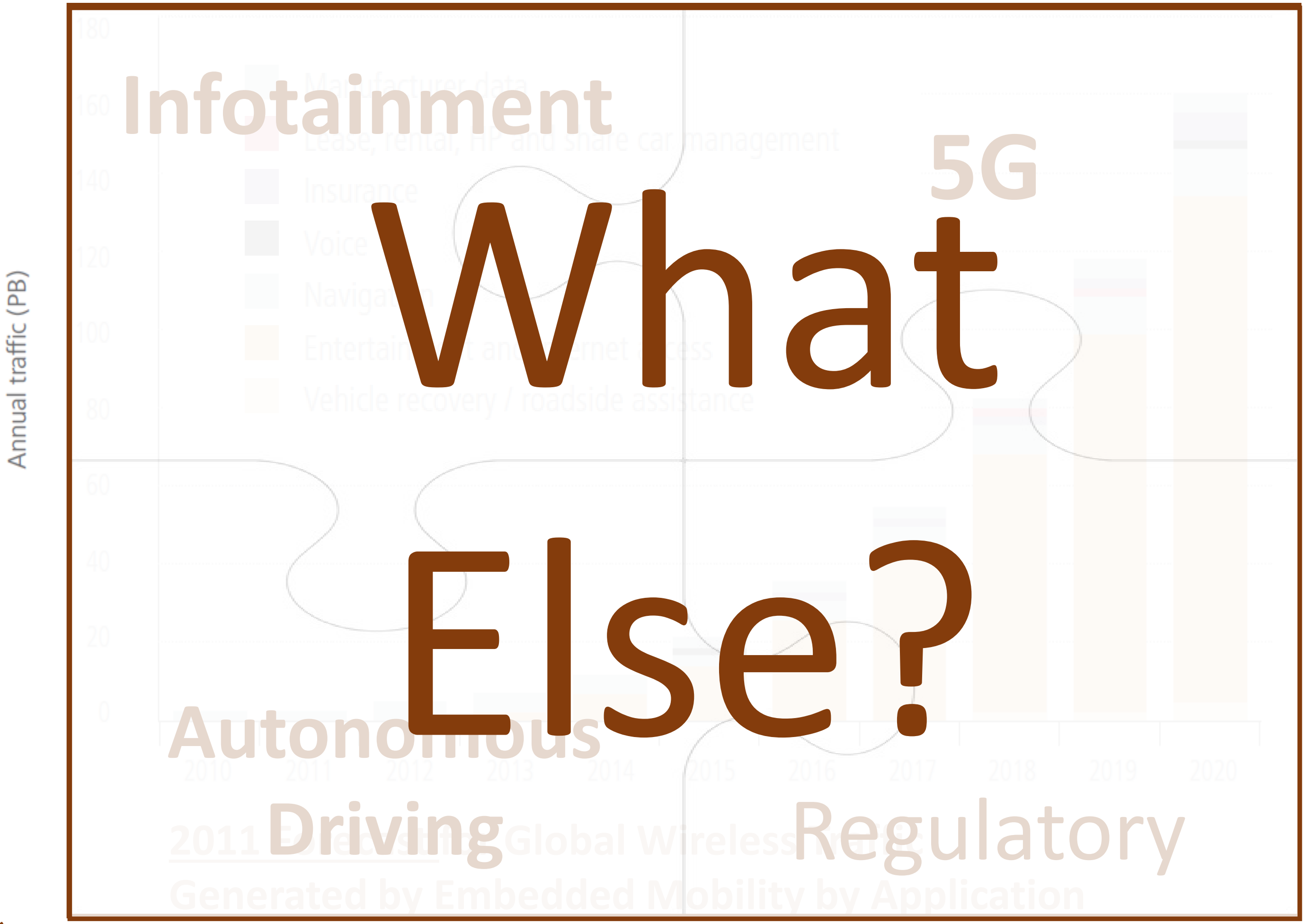


Connected Cars – Driving BW on Mobile Networks



2019- 117 Million Vehicles to be produced *

* CFI Multi-Gig Automotive Ethernet PHY, http://www.ieee802.org/3/cfi/1116_1/CFI_01_1116.pdf.



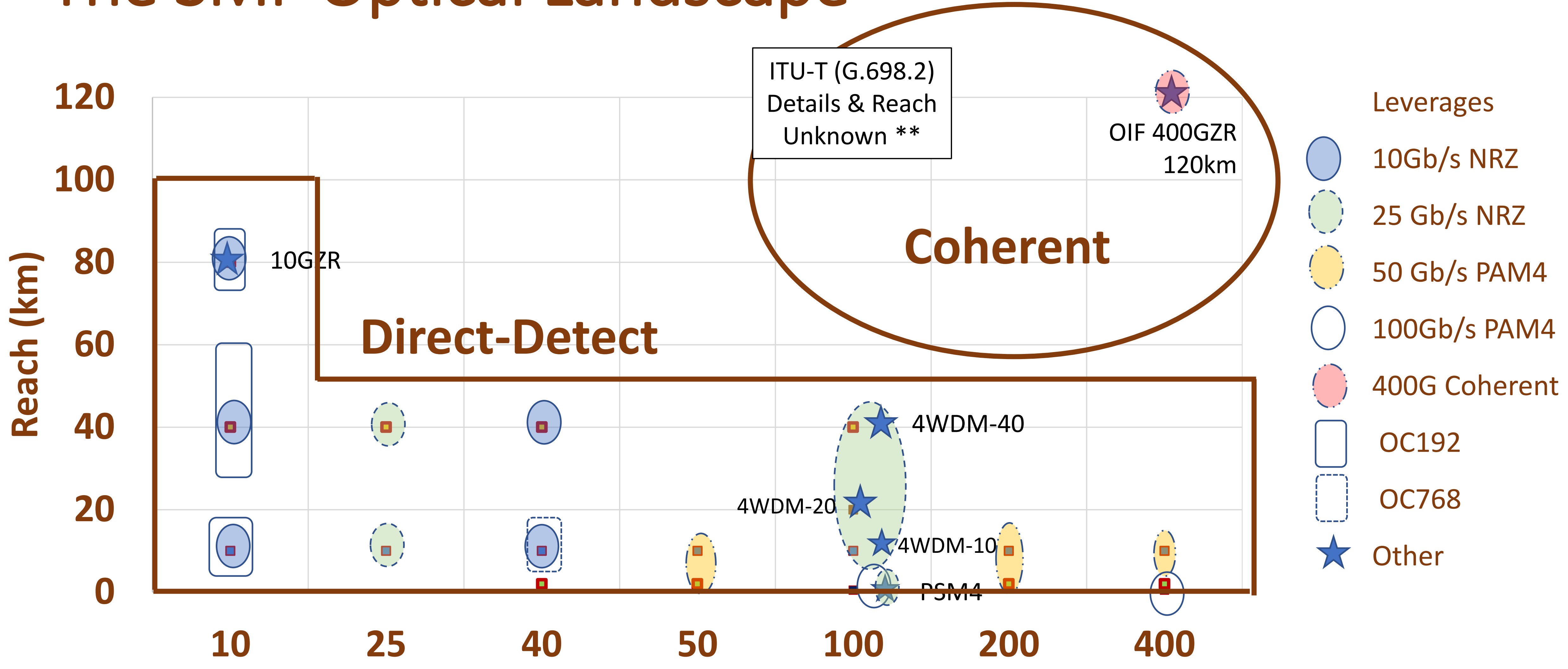
Source - GSMA, Connecting Cars: The Technology Roadmap, February 2013, https://www.gsma.com/iot/wp-content/uploads/2013/02/GSMA_mAutomotive_TechnologyRoadmap_v2.pdf

Summary

- 3 Million units (GbE to 100GbE) for 40km and beyond shipped annually
  - Not a data center application!
  - Bandwidth growth throughout EcoSystem
- “Geographically challenged” applications exist throughout Ecosystem
  - Internet Exchanges
  - Financial Industry
  - Mobile Backhaul
- China – Mobile Networks
  - Traffic in China alone exceeds other regions of the world
  - Consumer video driving application
- Emerging applications to drive future traffic over mobile networks

The Technical Aspect-
Beyond 10km Optical PHYs

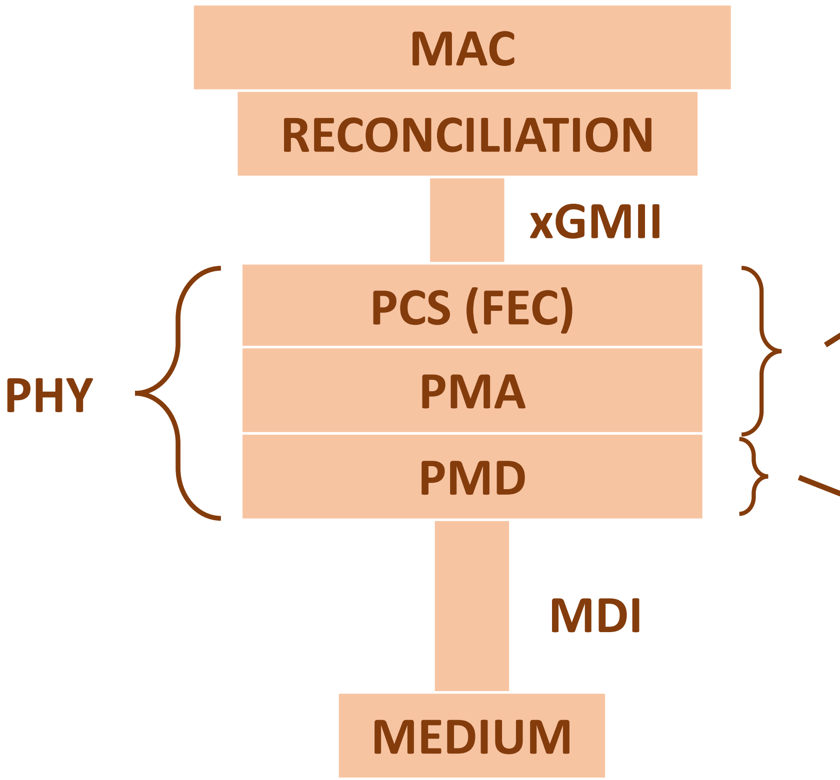
The SMF Optical Landscape *



* - Includes Standards and Efforts in development

** https://www.slideshare.net/ITU/itut-study-group-15-introduction.

An Ethernet Overview of the Problem

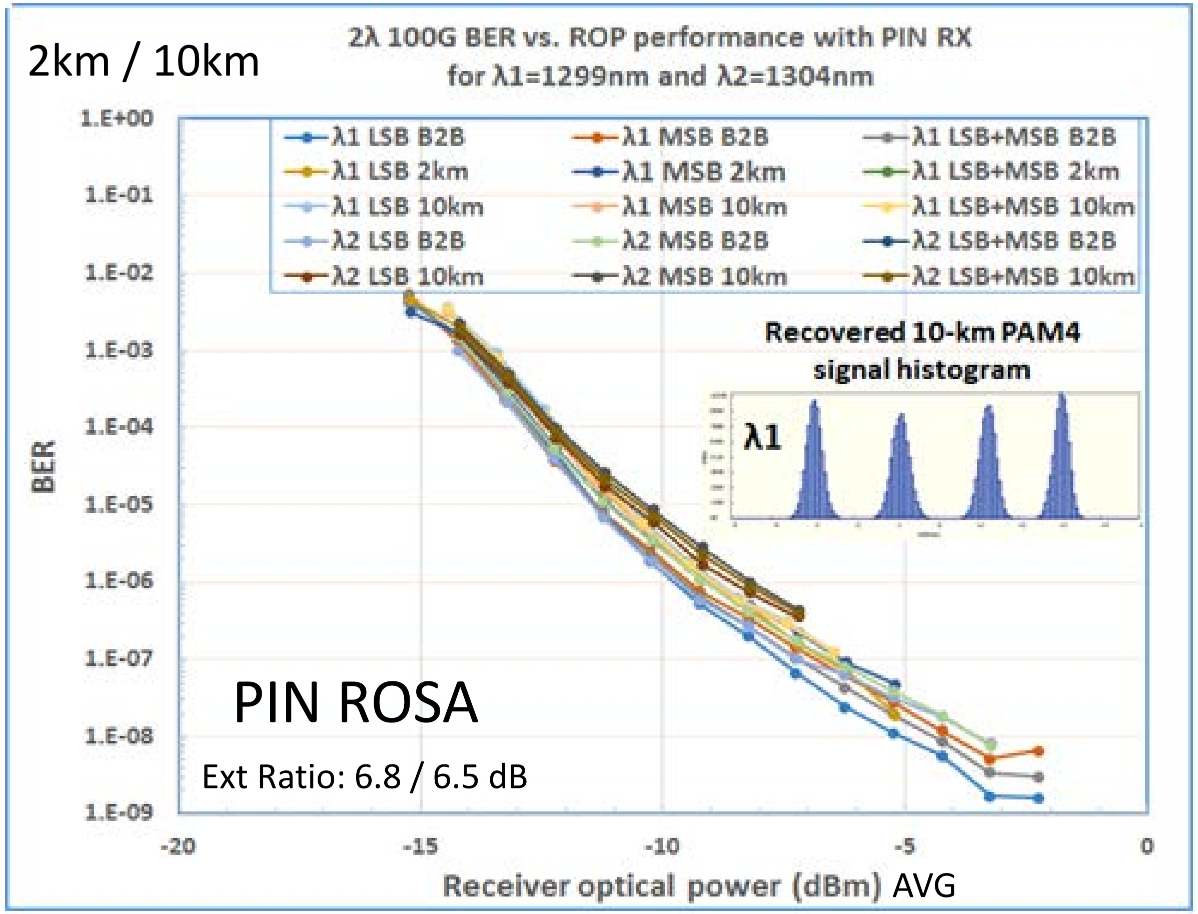


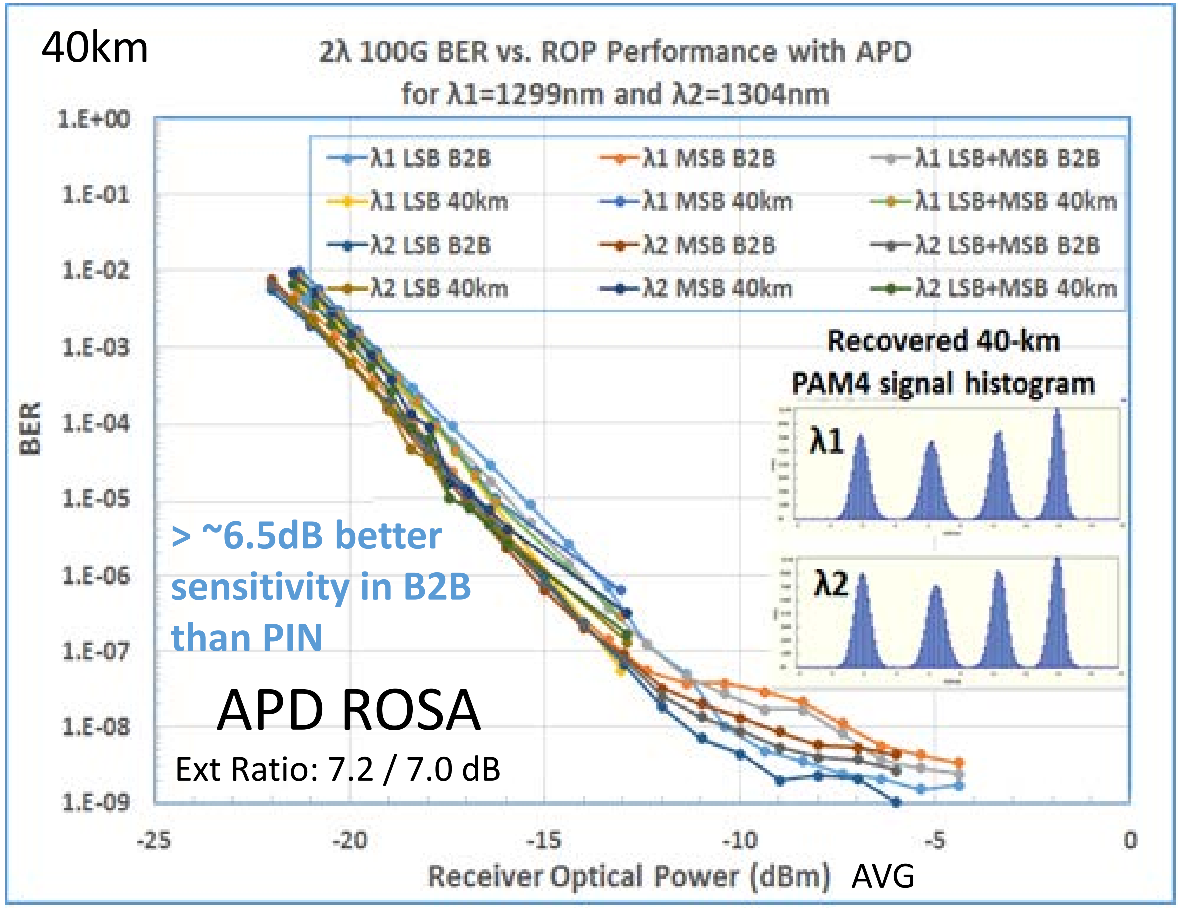
Approaches to Longer Reach

- Stronger FEC

- Improved Tx / Rx
- APD
- DSP
- Coherent
- Other (SOA, ….)

Impact of Use of APD (2λ @ 51.5625 Gb/s PAM4)

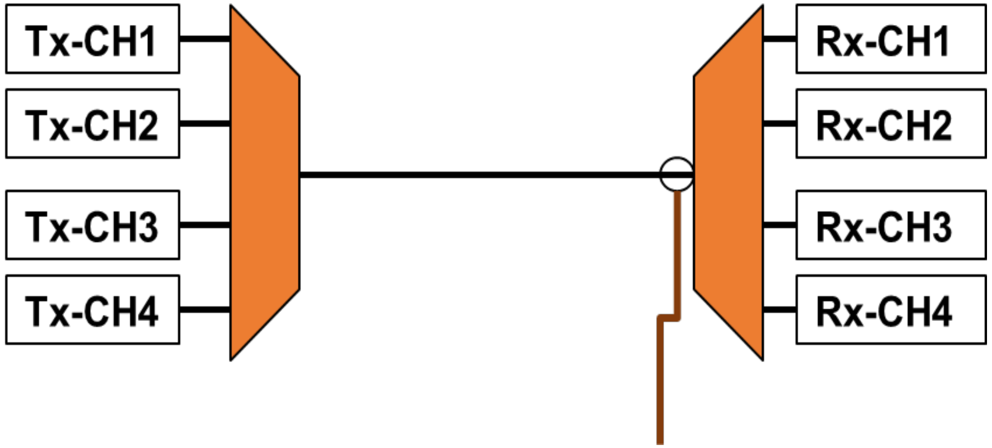




Data: PRBS31
Used actual chip implementation with real-time Rx DSP **with 10+ taps FFE** embedded inside the silicon

Source: Frank Chang, Inphi, "OFC 2016: Link Performance Investigation of Industry First 100G PAM4 IC Chipset with Real-time DSP for Data Center Connectivity "

4X50G PAM4 System Performance: BER



- Test method
  - Online test
  - All optical devices commercially available
  - Tx power (OMA) was adjusted to 5dBm
  - APD ROSA
  - Data Pattern – PRBS31

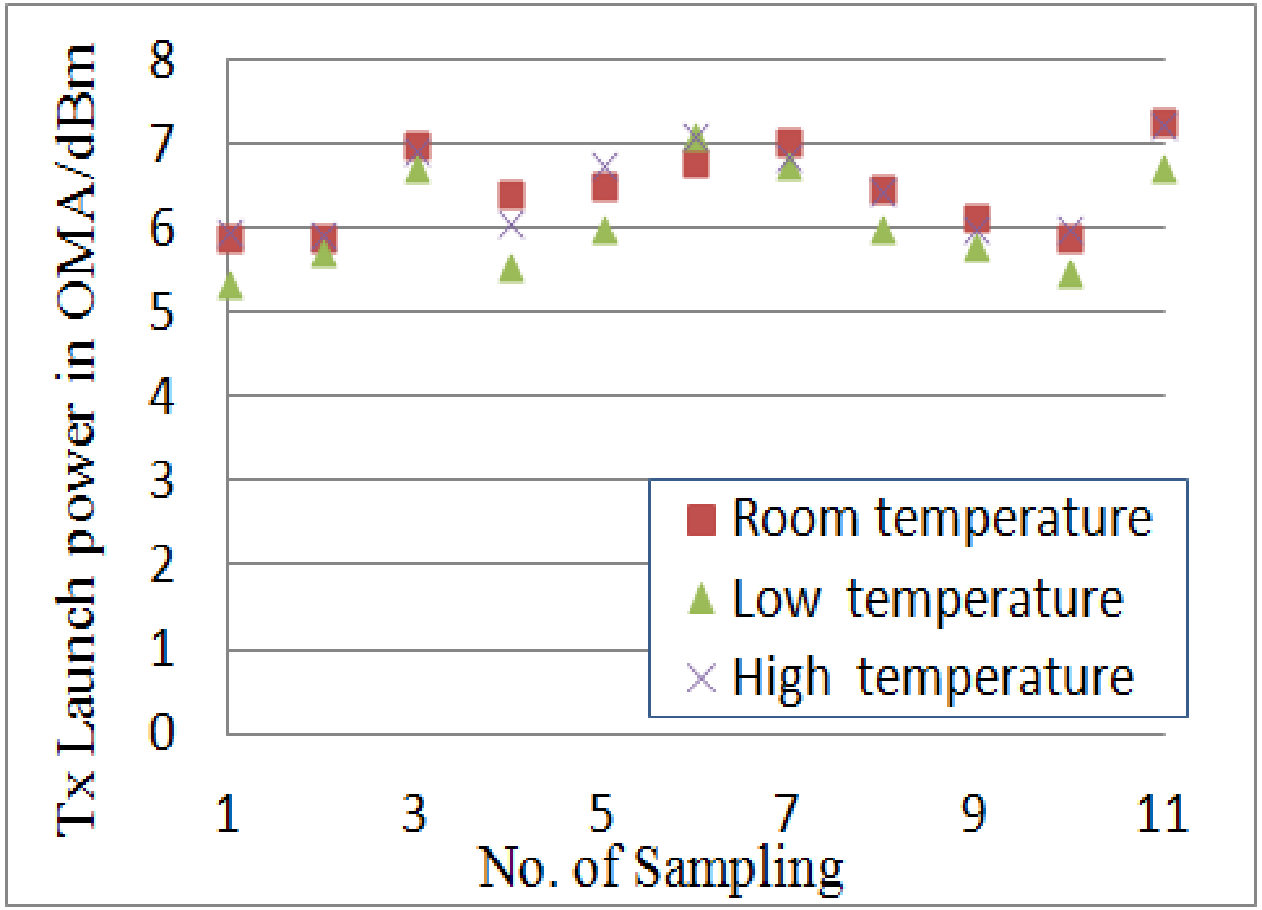
Lane	Center Wavelength
L0	1295.56nm
L1	1300.05nm
L2	1304.58nm
L3	1309.14nm

- Best/Worst case of BER test results (@ input to optical demux) of 11 random samples

	Tx Power (OMA dBm)	Rx Sensitivity (OMA dBm)	Budget (dB)	Temp (℃)	BER
Sample 1	5	-18.6	23.6	25	2.4E-4
Sample 2	5	-17.9	22.9	25	2.4E-4

Source: Xu Yu, Huawei

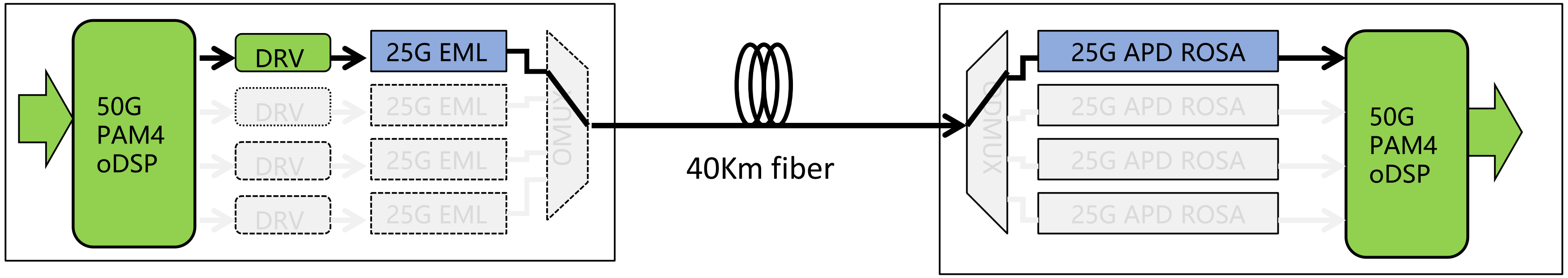
Tested Result of Transmitter Output Power



- **11 samples were tested in whole temperature range**
- **All of TX output power are higher than 5dBm, even under worst case.**
- **Note – Similar temperature testing of APD Receiver has not been tested yet.**
- **Temperatures**
  - **Low Temp – 0 C**
  - **Room Temp – 25 C**
  - **High Temp – 70 C**

Source: Xu Yu, Huawei

1X50G PAM4 System Performance: Dispersion Penalty



Suggested wavelength assignments:
Same wavelength as 50GBASE-LR

Lane	Center Wavelength	Wavelength Range
CH	1311nm	1304.5 to 1317.5 nm

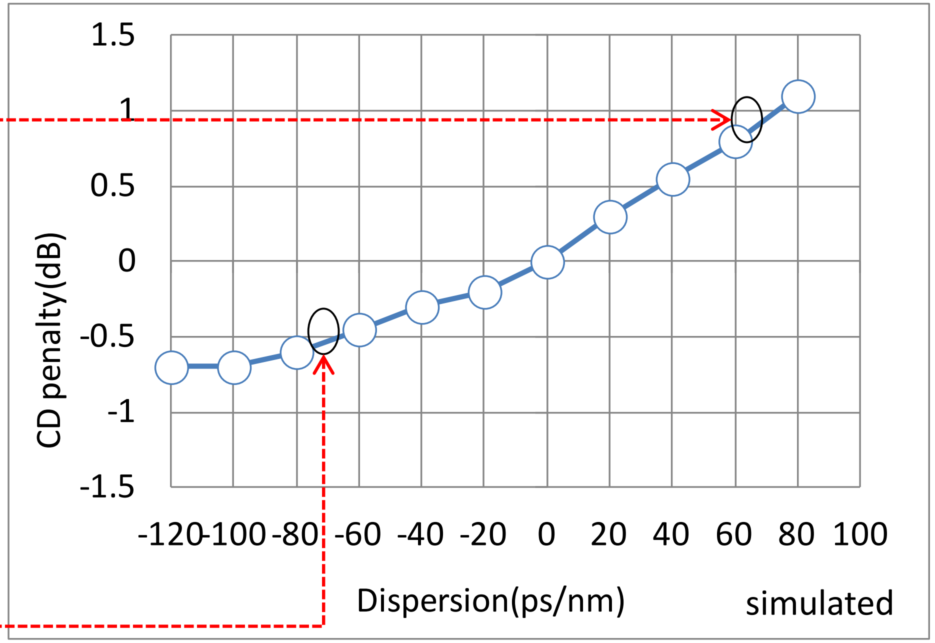
Worst case dispersion analysis:

$$Dispersion\ \ min.:\ 0.2325 * \lambda * \left[1 - \left(\frac{1324}{\lambda}\right)^4\right]$$

$$Dispersion\ \ max.: 0.2325 * \lambda * \left[1 - \left(\frac{1300}{\lambda}\right)^4\right]$$

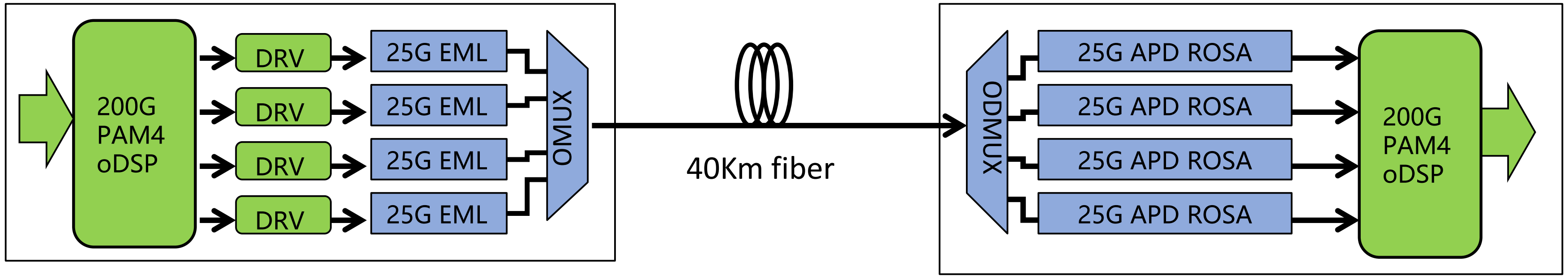
Item	Wavelength (nm)	Dispersion (λ_0=1300) ps/nm	Dispersion (λ_0=1324) ps/nm
1	1317.5	+63.81	--
2	1304.5	--	-74.18

Dispersion Penalty @2.4E-4



Source: Xu Yu, Huawei

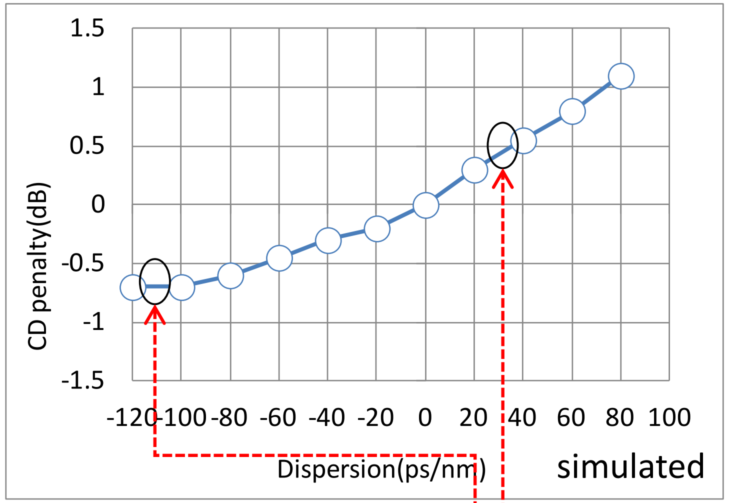
4X50G PAM4 System Performance: Dispersion Penalty



Suggested WDM assignments:
Same wavelength as 200GBASE-LR4

Lane	Center Frequency	Center Wavelength	Wavelength Range
L0	231.4THz	1295.56nm	1294.53~1296.59 nm
L1	230.6THz	1300.05nm	1299.02~1301.09nm
L2	229.8THz	1304.58nm	1303.54~1305.63nm
L3	229THz	1309.14nm	1308.09~1310.19nm

Dispersion Penalty @2.4E-4



simulated

Worst case dispersion analysis:

$$Dispersion\ min.:\ 0.2325 * \lambda * \left[1 - \left(\frac{1324}{\lambda}\right)^4\right]$$

$$Dispersion\ max.: 0.2325 * \lambda * \left[1 - \left(\frac{1300}{\lambda}\right)^4\right]$$

Lane	Wavelength (nm)	Dispersion (λ_0=1300) ps/nm	Dispersion (λ_0=1324) ps/nm
L0	1294.53	--	-113.43
L3	1310.19	+37.47	--

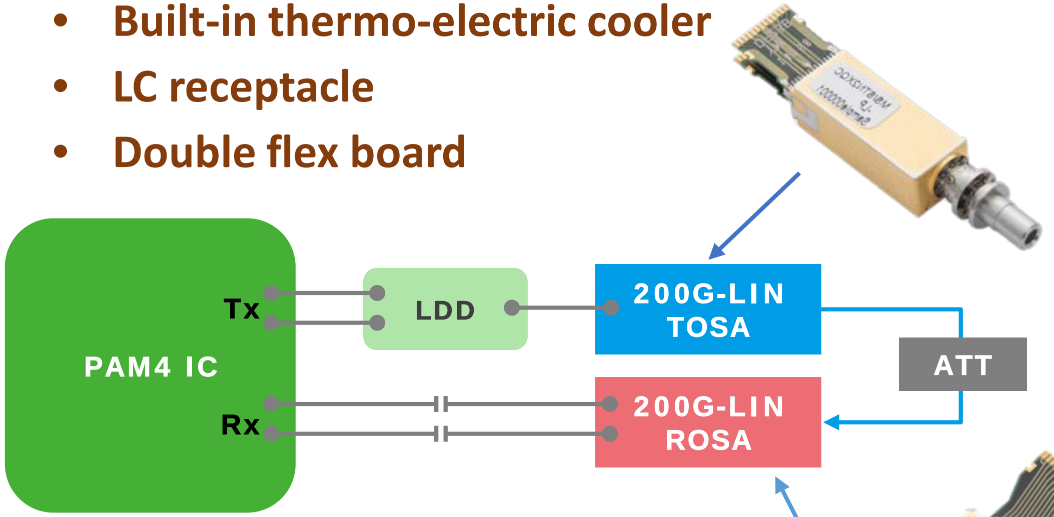
Source: Xu Yu, Huawei

200Gb/s PAM4 Test Results (50Gb/s x 4λ EML Linear TOSA & APD Linear ROSA)

Test Setup

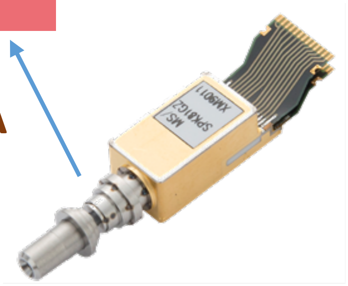
4λ integrated linear TOSA

- **LAN-grid 4x50G PAM4 EML**
- **Built-in thermo-electric cooler**
- **LC receptacle**
- **Double flex board**



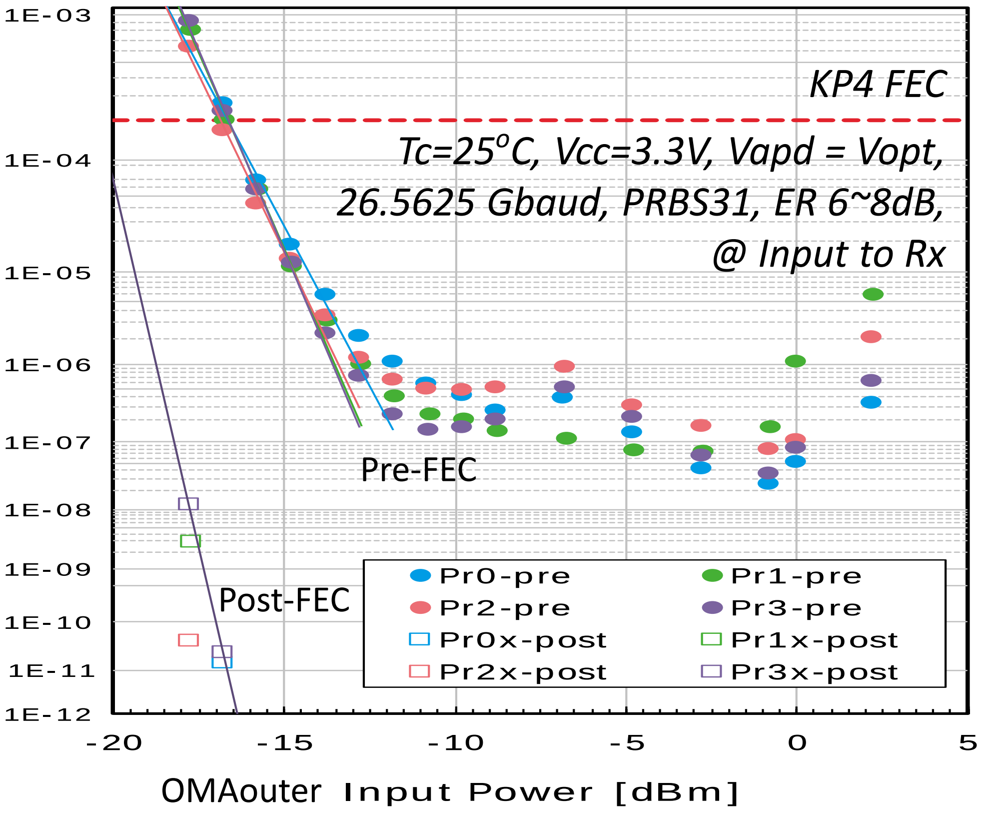
4λ integrated linear ROSA

- **4x50G PAM4 InP APD**
- **Quad linear TIA**
- **LC receptacle**
- **Double flex board**



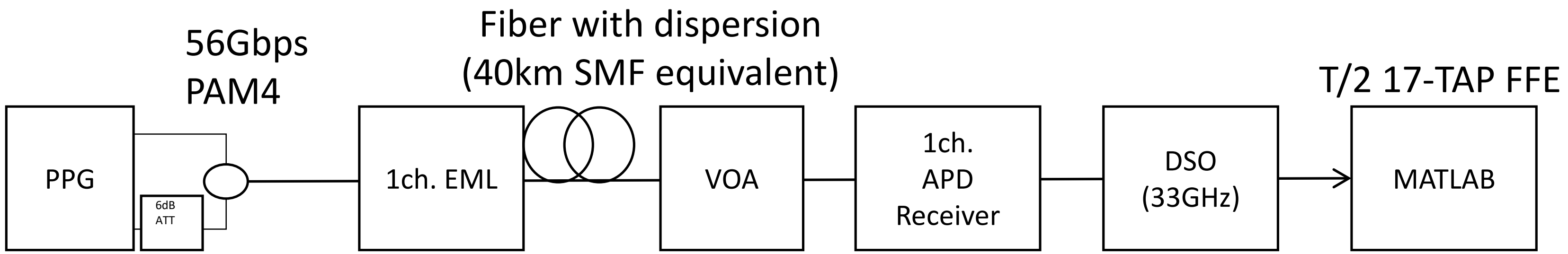
Source: Kenneth Jackson, Sumitomo Electric Device Innovations, USA

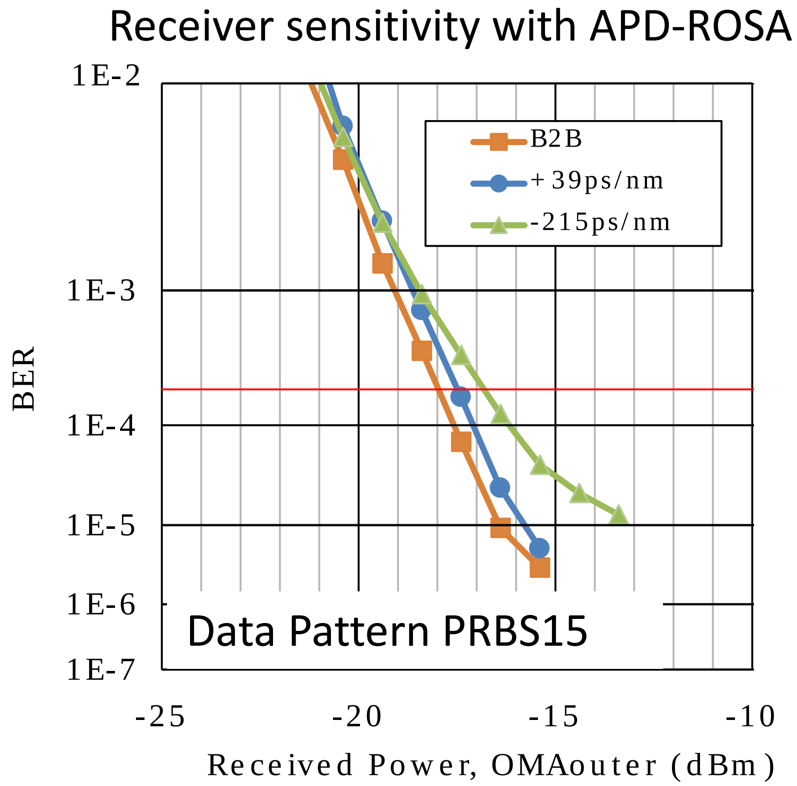
Sensitivity (Back-to-back)



"Pre" BER data-points for power levels ≥ -15dBm correctable to at least 30 sec error free.

Receiver sensitivity with APD ROSA





56G PAM4 reach extension is achieved.

APD receiver can achieve rec. sensitivity of

-16.7 dBm for the worst case dispersion(neg.)*

-18.0 dBm for the worst case dispersion (B2B)

*** assumed 8-lane LAN-WDM over SMF**

**** Better than typical APD**

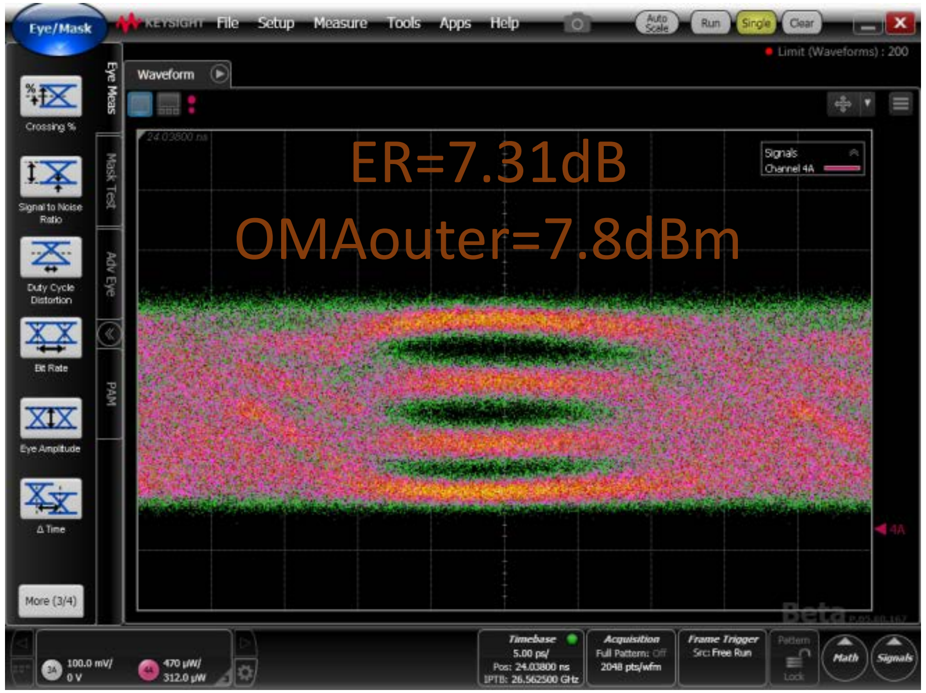
Assuming KP4 FEC but still 56Gpps can accommodate overhead associated with stronger FEC.

Source: Yoshiaki Sone, NTT

Link budget example with High-power EML

Evaluation result using high power EML and APD- ROSA
Link- budget=24.9dB(1ch B2B, KP4 FEC limit)

PAM4 tx eye with high power EML



7.8dBm
(EML)

24.9dB* (1ch-B2B)

20.9dB** (4ch-B2B)

18.9dB** (8ch-B2B)

-17.16dBm
(APD=ROSA)

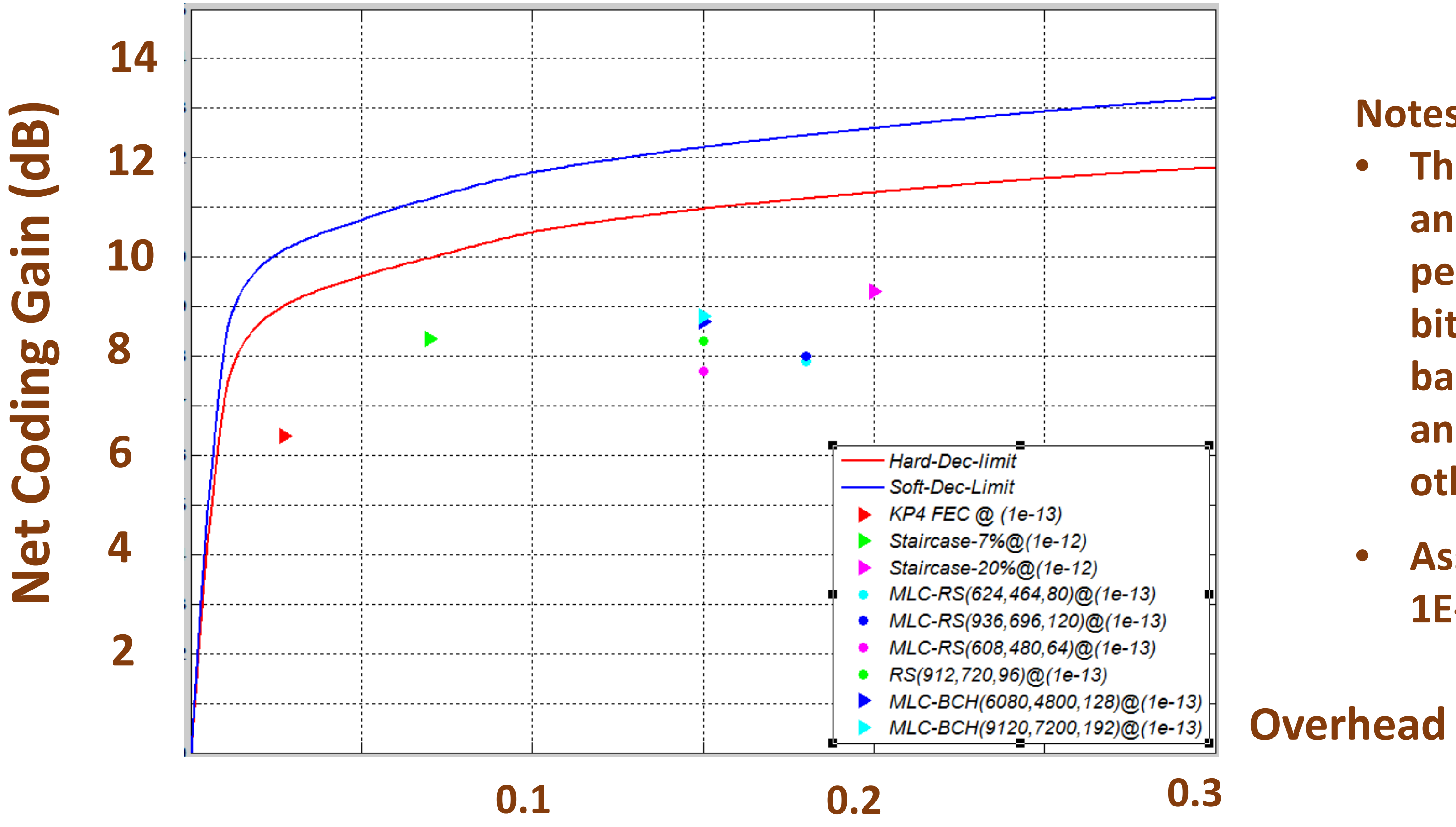
* WDM mux/demux loss is not included

** 4:1 Ratio 2 dB, 8:1 Ration 3 dB mux/demux loss (see http://www.ieee802.org/3/bs/public/adhoc/smf/14_09_30/cole_01_0914_smf.pdf)

Source: Yoshiaki Sone, NTT

Use of Stronger FEC

Several Potential HD-FECs can help to achieve beyond 10km 400GbE
RS-FEC, BCH-FEC, MLC-FEC or Staircase FEC. ([wang_ecdc_01_0316])

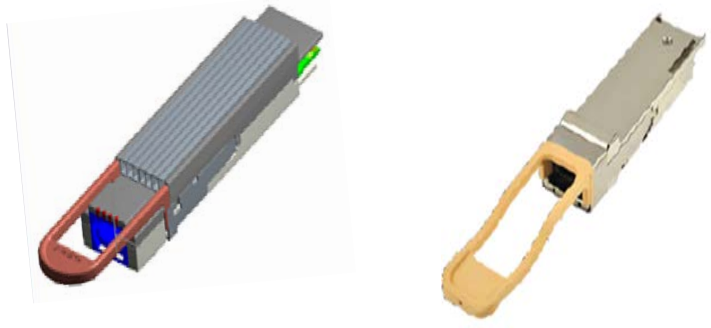


Notes –

- **This is a theoretical analysis that assumes penalty for increased bit rate is just the noise bandwidth increase and does not include other penalties.**
- **Assumes post BER @ 1E-13 objective**

The OIF 400ZR Project

- Implementation agreement (IA) for pluggable digital coherent optical (DCO) modules
  - Amplified short-reach DWDM applications with distances up to 120 km
  - Passive single channel ZR (80km)
- Single-carrier 400 G, coherent detection and advanced DSP / FEC algorithms.
- Operates as a 400 GbE PMD compatible with 400G-AUI.
- Other formats could be considered in the project as well.
- Supporters from more than 34 companies, including end users, system and component suppliers. Unanimous support for start of project



Form Factors

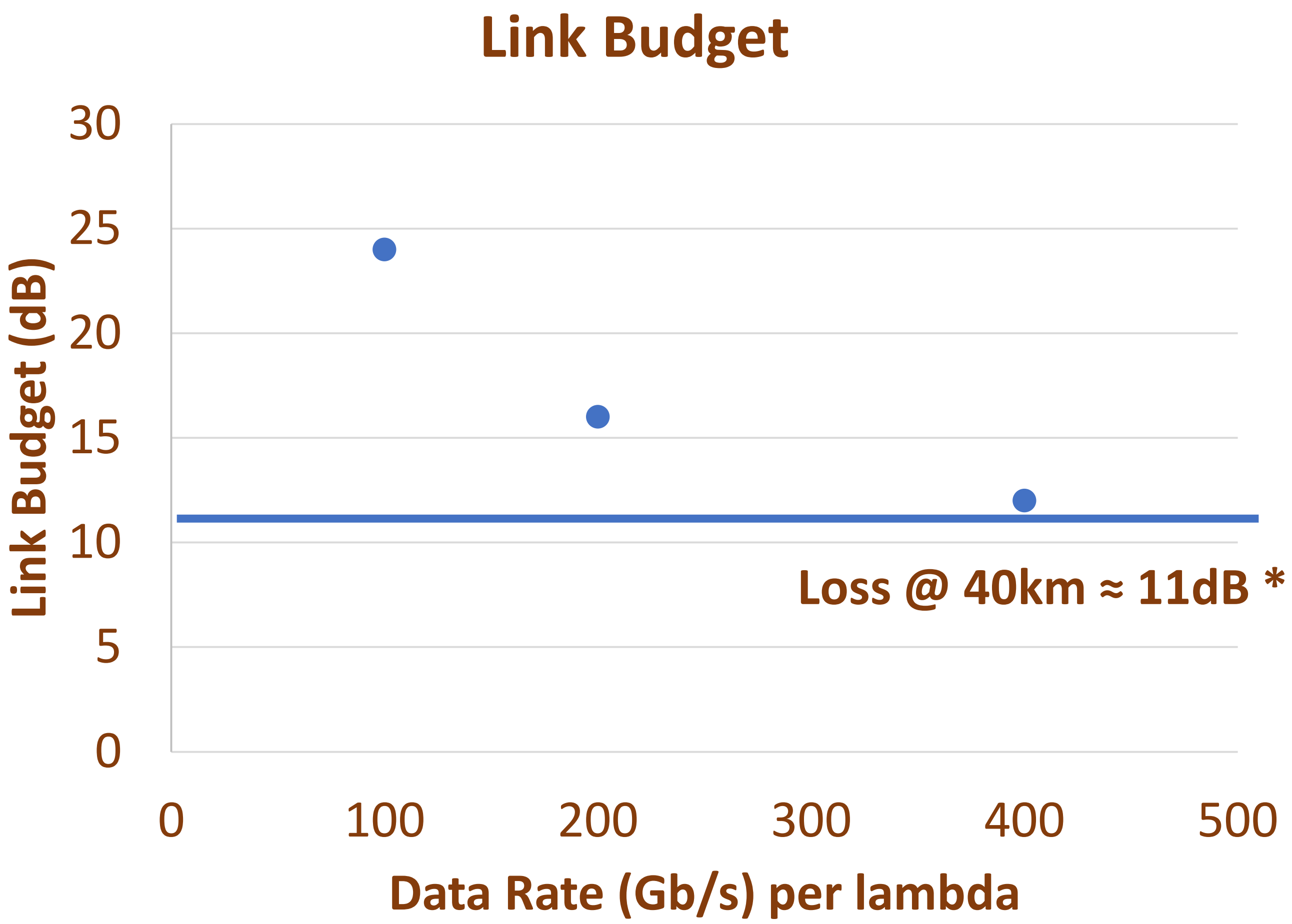
Targeting coherent optics in client pluggable form factors <15W

Assumes tunable λ not required for this application

Source: Tom Williams, Acacia

Source: OIF Liaison to IEEE 802.3, Nov 7, 2016:
http://www.ieee802.org/3/minutes/nov16/incoming/OIF_to_IEEE_802d3_Nov_2016.pdf

Targeting 40km with Coherent Technology



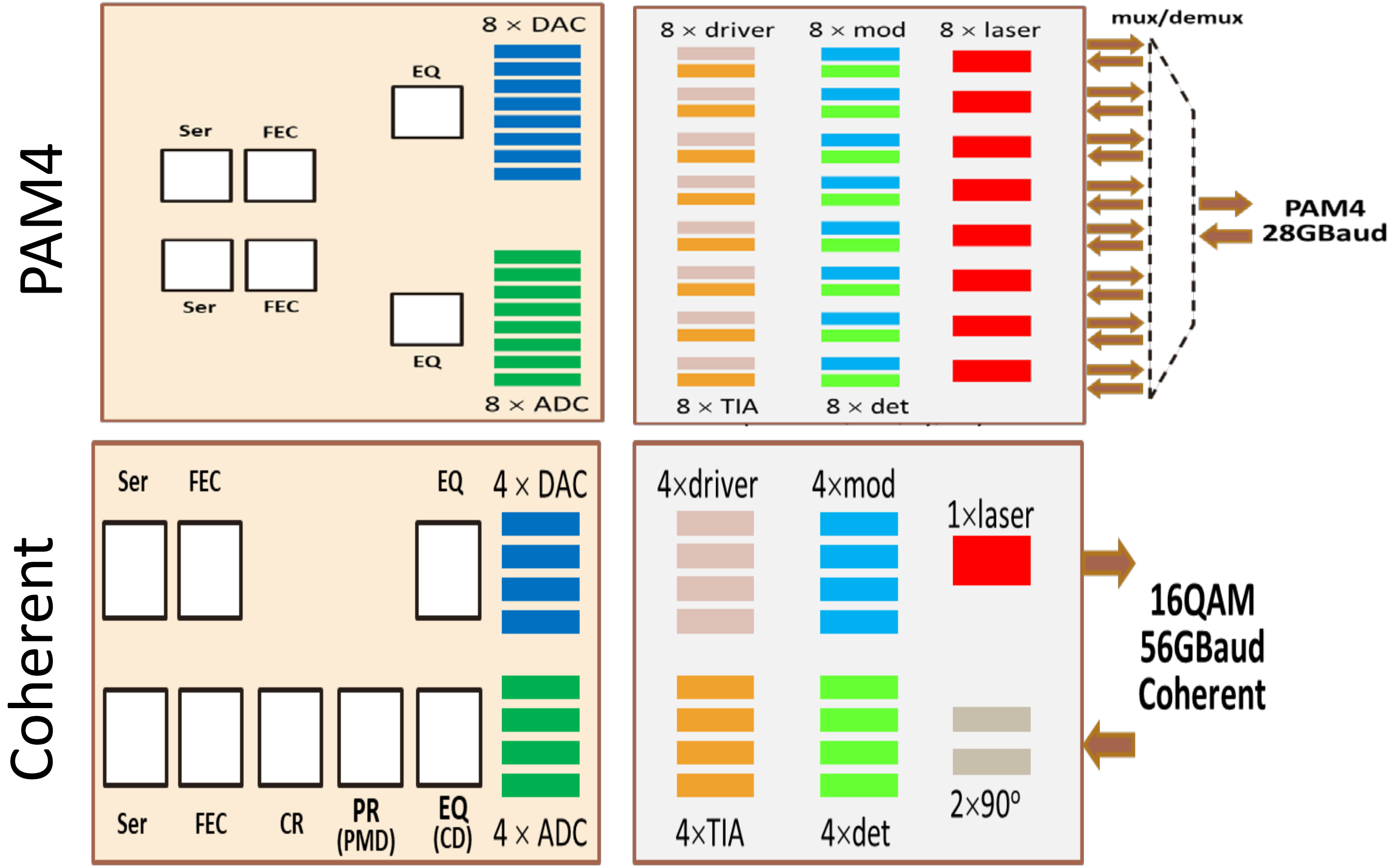
Assumptions

- Modulation Format
  - 100G – QPSK @ ~30Gbaud
  - 200G – 16QAM @ ~30Gbaud
  - 400G – 16QAM @ ~60Gbaud
- Tx and Rx power levels achievable with high yield and multiple optical technologies
- Note – Longer reach, ie. higher link budgets, can be supported by transmit SOA/EDFA or with additional amplification

* - http://www.ieee802.org/3/ba/public/tools/Fibre_characteristics_V_3_0.xls

Source: Tom Williams, Acacia

Implementation Cost Considerations



Implementation costs need to be studied –

- **Inclusion of components**
- **Number of components**
- **Operation rate of components**
- **Specifications of components**

Source: Tom Williams, Acacia

Technical Feasibility of Beyond 10km Optical PHYs

- **Growing evidence of different ways to support reaches beyond 10km for 50GbE, 200GbE, 400GbE**
  - **PAM4 (Direct Detect) test data for 40km provided**
    - **Higher Power EML Transmitters**
    - **APDs**
    - **Advanced DSP**
    - **FEC**
  - **Coherent Optics**
    - **Shipping today**
    - **Industry development efforts that may be leveraged.**
      - **ITU-T (ITU-T G.698.2)**
      - **OIF 400GZR (120km)**
- **Real challenge – determining the right solution for the right reach / rate!**

Why Now?

Why Now?

- Applications for Beyond 10km Optical PHYs
  - Everywhere - ≈3M units shipped annually addressing 40+km
  - Example application space– Mobile Backhaul Networks
    - Mobile Networks in China illustrate the impact of consumer video
    - Other examples of "geographically challenged" reaches highlighted– Financial, Metro
    - Emerging future bandwidth growth driver- Automotive
  - Not same volumes as Data Center – but relevant to overall EcoSystem
- Traffic is growing everywhere
  - More users
  - More ways to access the internet faster
  - Higher bandwidth content
  - New applications enabled
  - And it goes on
- There are no optical Ethernet solutions for Beyond 10km for 50GbE, 200GbE, and 400GbE
- Time is not on our side…

Contributors

John D’Ambrosia, Futurewei, Subsidiary of Huawei

Thanks to the following individuals for their input or slides -

- **Pete Anslow, Ciena**
- **Andrew Bach, Independent**
- **Steve Carlson, High Speed Design**
- **Frank Chang, Inphi**
- **Weiqiang Cheng, China Mobile**
- **Lu Huang, China Mobile**
- **Alexander Ilin, MSK-IX**
- **Kenneth Jackson, Sumitomo Electric Device Innovations, USA**
- **David Lewis, Lumentum**
- **Dale Murray, LightCounting**
- **Gary Nicholl, Cisco**
- **Yoshiaki Sone, NTT**
- **Xinyuan Wang Huawei**
- **Tom Williams, Acacia**
- **Alexander Umnov, Corning**
- **Xu Yu, Huawei**
- **Wenyu Zhao, CAICT**

Thanks to IEEE 802.3 New Ethernet Applications Ad hoc for feedback

Supporters

- Justin Abbott, Lumentum
- Thananya Baldwin, Ixia
- Vipul Bhatt, Finisar
- Martin Bouda, Fujitsu Laboratories of America
- Patricia Bower, Socionext Europe GmbH
- Ralf-Peter Braun, DT
- Paul Brooks, Viavi Solutions
- Matt Brown, MACOM
- Li Cao, Accelink
- Steve Carlson, High Speed Design
- Derek Cassidy, ICRG
- David Chalupsky, Intel
- Frank Chang, Inphi
- Xin Chang, Huawei
- David Chen, AOI
- James Chien, ZTE
- Chris Cole, Finisar
- Doug Coleman, Corning
- John D'Ambrosia, Futurewei, Subsidiary of Huawei
- Curtis Donahue, UNH-IOL
- Mike Dudek, Cavium
- David Estes, Spirent
- Kazuhisa Furusawa, NTT
- Ali Ghiasi, Ghiasi Quantum
- Zhigang Gong, O-Net Communications
- Mark Gustlin, Xilinx
- Ruibo Han, China Mobile
- Akinori Hayakawa, Fujitsu Laboratories
- Riu Hirai, Hitachi Ltd.
- Lu Huang, China Mobile
- Mengyuan Huang, Sifotonics
- Jeff Hutchins, Ranovus
- Jonathan Ingham, Foxconn Interconnect Technology
- Kazuhiko Ishibe Anritsu
- Hideki Isono, Fujitsu Optical Components
- Tom Issenhuth, Issenhuth Consulting / Huawei
- Ken Jackson, Sumitomo Electric Device Innovators, USA
- John Johnson, Broadcom
- Yasuaki Kawatsu Appresia systems
- Nobuhiko Kikuchi, Hitachi Ltd.
- Mark Kimber, Semtech
- Jonathan King, Finisar
- Curtis Knittle, Cable Labs
- Jeff Lapak, UNH-IOL
- Greg Lecheminant, KeySight
- Hanan Leizerovich, MultiPhy
- David Lewis, Lumentum
- Jon Lewis, Dell EMC
- Junjie Li, China Telecom
- Mike Li, Intel
- Robert Lingle, OFS
- Hai-Feng Liu, Jntel
- Samuel Liu, Nokia
- Scott Kipp, Brocade
- Jeff Maki, Juniper
- David Malicoat, SENKO Advanced Components
- Tom McDermott, Fujitsu Network Communications
- Greg McSorley, Amphenol
- Rich Mellitz, Samtec
- Christophe Metivier, Arista
- Dale Murray, LightCounting
- Gary Nicholl, Cisco
- Paul Nikolich, Independent
- Mark Nowell, Cisco
- David Ofelt, Juniper
- Tom Palkert, Molex
- Earl Parsons, Commscope
- Vasu Parthasarathy, Broadcom
- Gerry Pepper, Ixia
- David Piehler, Dell EMC
- Dino Pozzebon, Microsemi
- Rick Rabinovich, IXIA
- Stefan Rochus, Broadcom
- Salvatore Rotolo, ST Microelectronics
- Ed Sayre, Teraspeed, a Division of Samtec
- Mizuki Shirao, Mitsubishi
- Kapil Shrikhande, Innovium
- Scott Sommers, Molex
- Yoshiaki Sone, NTT
- Ted Sprague, Infinera
- Rob Stone, Broadcom
- Phil Sun, Credo Semiconductor
- Steve Swanson, Corning
- Akio Tajima NEC
- Tomoo Takahara, Fujitsu Laboratories
- Kohichi Tamura, Oclaro
- Brian Teipen, ADVA Optical
- Nathan Tracy, TE Connectivity
- Matt Traverso, Cisco
- David Tremblay, HPE
- Ed Ulrichs, Source Photonics
- Alexander Umnov, Corning
- Haijun Wang, China Unicom
- Xinyuan Wang, Huawei
- Winston Way, NeoPhotonics
- Markus Weber, Acacia
- Brian Welch, Luxtera
- Tom Williams, Acacia
- Qing Xu, Belden
- Yu Xu, Huawei
- Ryan Yu, Oplink Communication
- Wenyu Zhao, CAICT
- Huanlin Zhang, AOI
- George Zimmerman, CME Consulting
- Pavel Zivny Tektronix

Straw Polls

Call-For-Interest

- **Should a Study Group be formed to consider Beyond 10km Optical PHYs for 50GbE, 200GbE, and 400GbE?**

Y: 82 N: 0 A: 16

Room Count: 103

Participation

- **I would participate in the "Beyond 10km Optical PHYs" Study Group in IEEE 802.3.**

  Tally: 57

- **My company would support participation in the "Beyond 10km Optical PHYs" Study Group in IEEE 802.3**

  Tally: 39

Future Work

- Ask 802.3 on Thursday
  - Form Beyond 10km Optical PHYs Study Group
  - Approve Liaisons regarding SG formation & requesting status update, pending all approvals
    - OIF
    - ITU-T
- If approved, on Friday
  - Request 802 EC form Beyond 10km Optical PHYs Study Group
  - First Beyond 10km Optical PHYs SG meeting, week of Sept 2017 IEEE 802.3 Interim
  - Teleconference Calls to be scheduled
  - Liaisons announcing SG formation & requesting status update
    - OIF
    - ITU-T