

IEEE 802.3 NEA Ad hoc 14 Sept 2020

IEEE 802.3 Call for Interest
CFI Consensus Presentation
Draft Development

“Beyond 400 Gb/s Ethernet”

John D'Ambrosia
Futurewei Technologies
U.S. Subsidiary of Huawei



OBJECTIVE FOR THE MEETING

- To measure the interest in starting a study group to address “Beyond 400 Gb/s Ethernet”
- 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 on the call may speak / vote
- RESPECT... give it, get it

Contributors

- John D'Ambrosia, Futurewei, U.S. Subsidiary of Huawei
- Matt Brown, Huawei Canada
- Joel Goergen, Cisco
- Mark Gustlin, Cisco
- Cedric Lam, Google
- David Piehler, Dell-EMC
- Ted Sprague, Infinera
- Rob Stone, Facebook
- Jim Theodoras, HG Genuine
- Nathan Tracy, TE Connectivity
- Tedros Tsegaye, Innolight
- Xinyuan Wang, Huawei

➤ Also

- IEEE 802.3 2020 Ethernet Bandwidth Assessment
- IEEE 802.3 NEA Ad hoc

Today's Panel

- To be identified

DRAFT DEVELOPMENT - FINAL VERSION PENDING

AGENDA

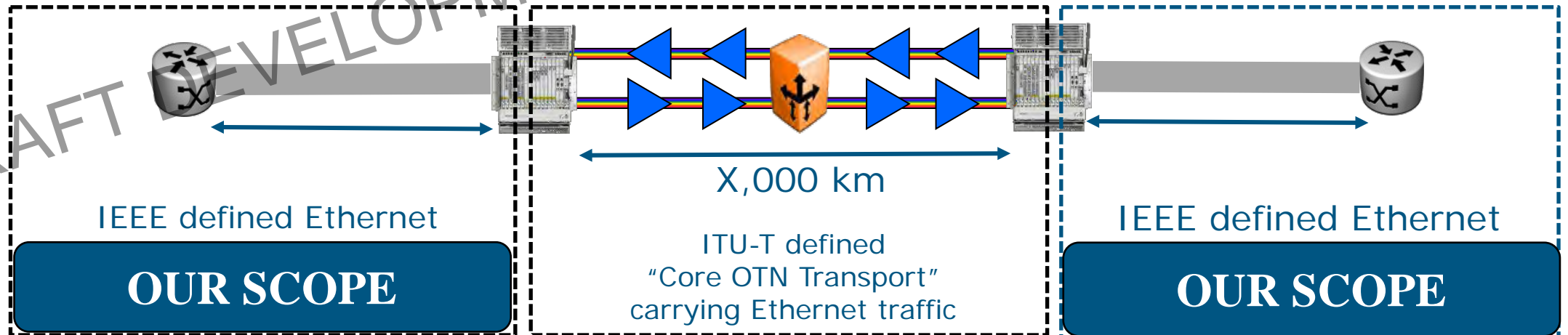
- **Introduction**
- **Presentations**
 - **Market Pressures for Beyond 400 GbE**
 - **The Technical Roadmap to Beyond 400 GbE**
 - **Beyond 400 GbE - Why Now?**
- **Straw Polls**
- **Future Work**

THE SCOPE OF ETHERNET TODAY

Scenario #1



Scenario #2



Potential for Technology Reuse

Reuse of signaling rate technologies developed for higher Ethernet rates enables existing lower speed Ethernet rate specifications (AUI, -KR, -CR, -SR, -DR, -LR, -ER)



Image courtesy of David Piehler, Dell-EMC

- 32 400 Gb/s capacity ports
- Can be configured to support 32 400 GbE ports
- Can be configured to support 128 100 GbE ports



“It has been my experience at Google that we have used optical and cu modules to support different configurations of a given port, including applications that require the maximum capacity of the single port.”

Cedric Lam, Google

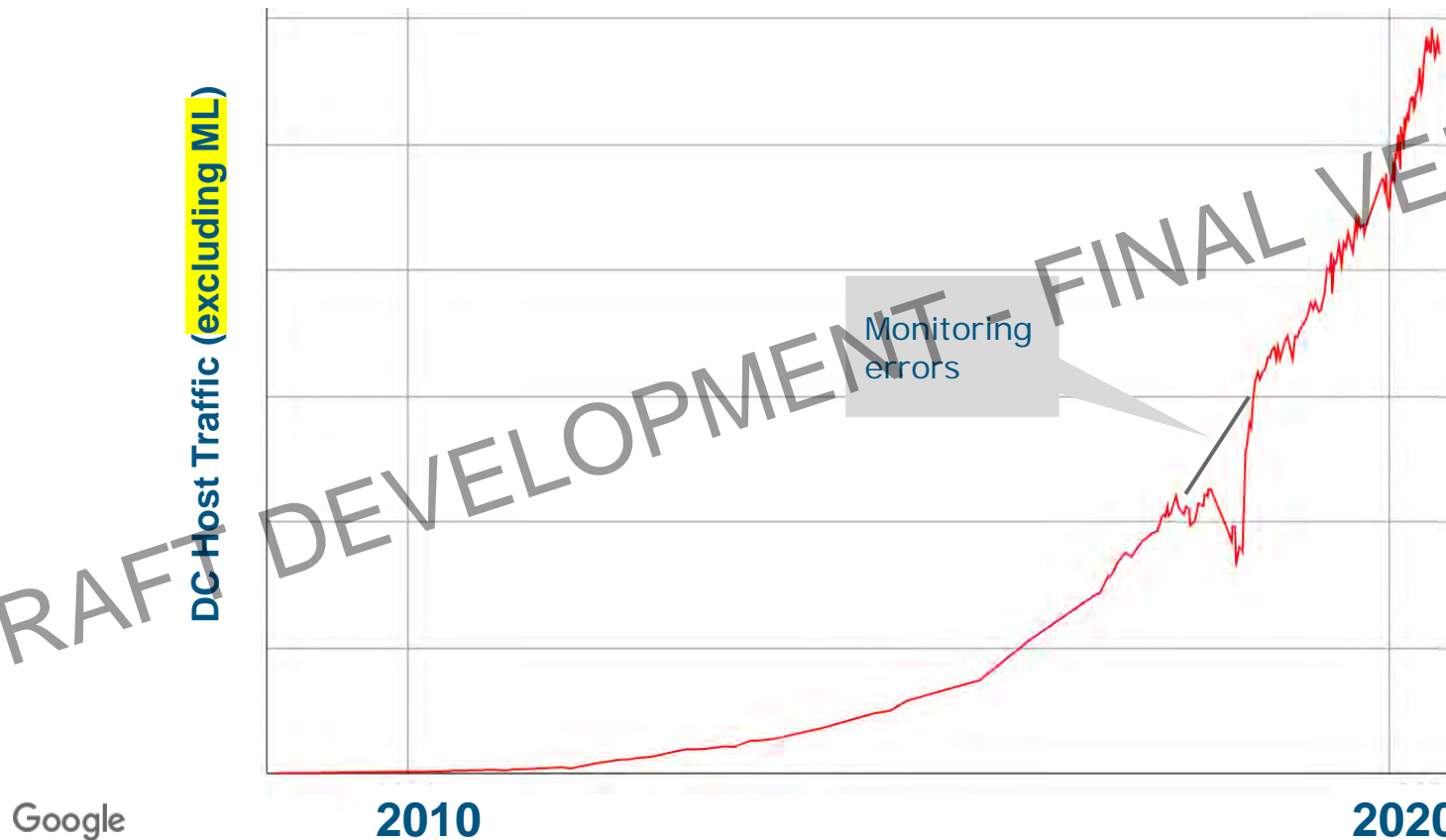
MARKET PRESSURES FOR BEYOND 400 GbE

DRAFT DEVELOPMENT - FINAL VERSION PENDING



DATA CENTERS CONTINUE AS A PRIMARY DRIVER

DC Traffic Continues to Grow Rapidly (Regular Servers)



Google

Courtesy - Cedric Lam, Google

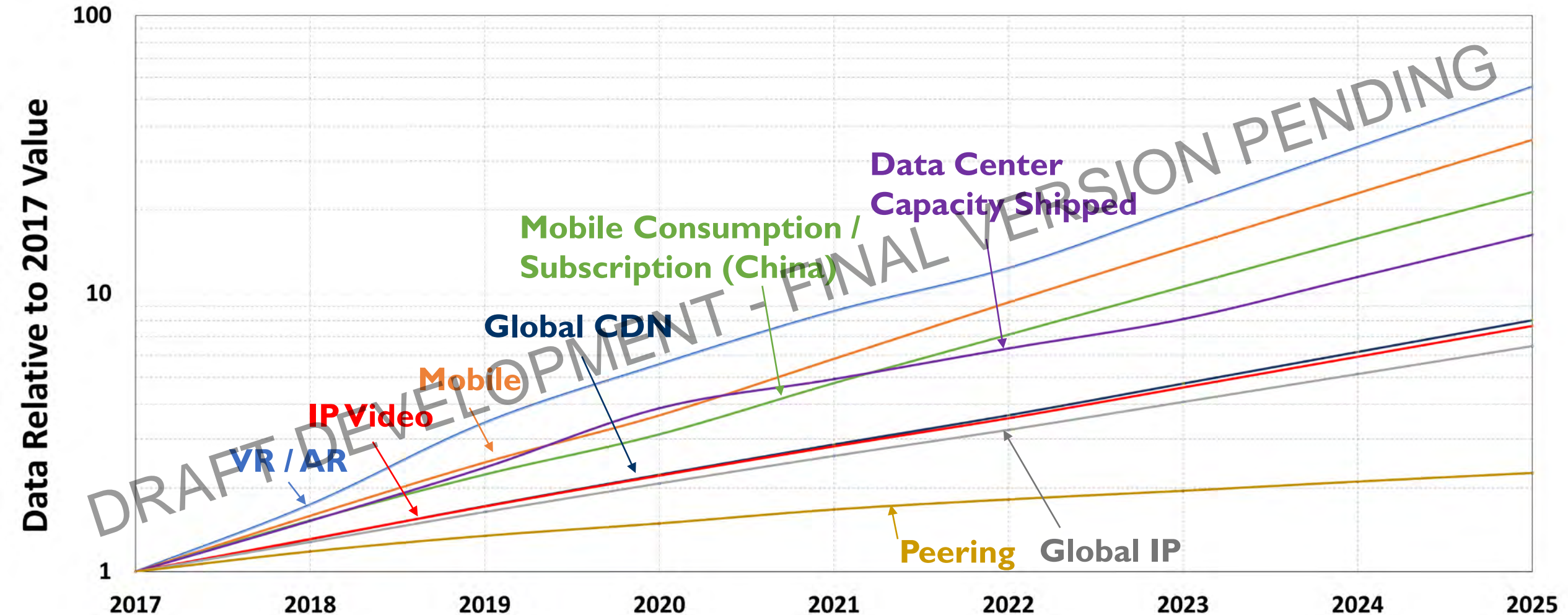
THE SONG REMAINS THE SAME

- 2020 Ethernet Bandwidth Assessment (BWA) documented latest analysis of industry bandwidth needs and driving factors

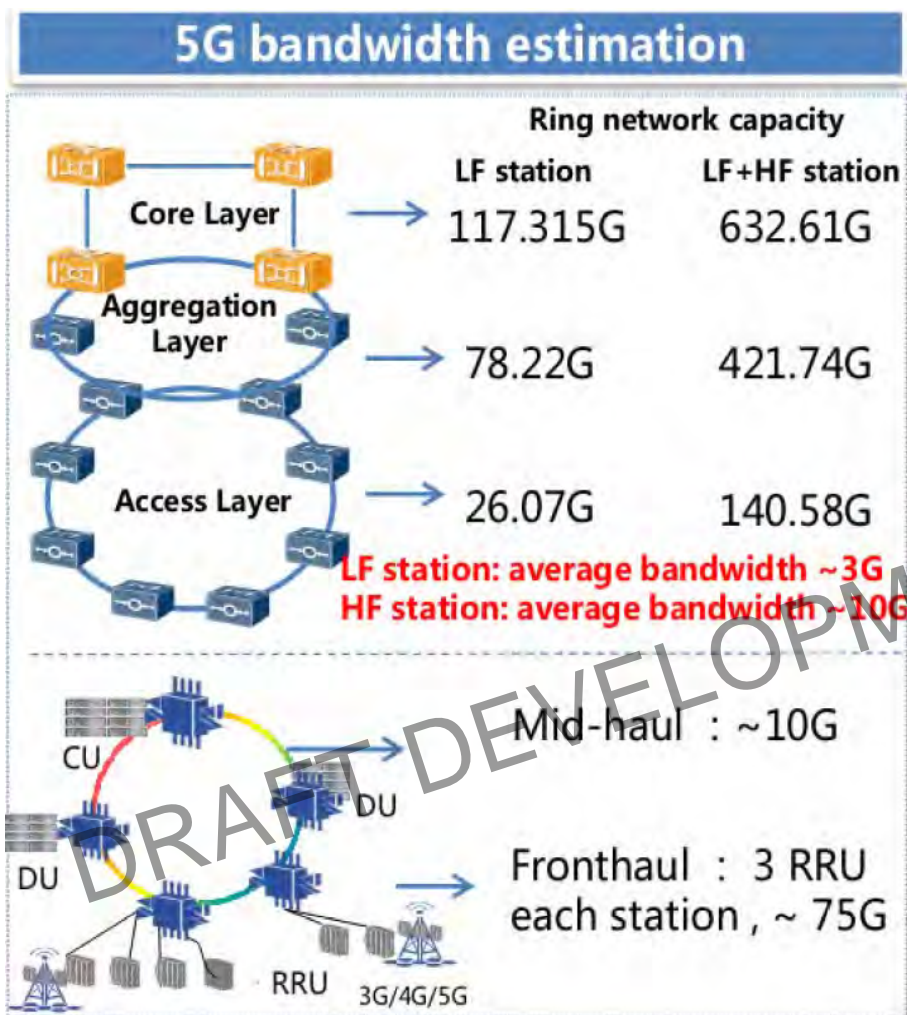
Increased # of users x Increased access methods and rates x Increased services = Bandwidth Explosion

- 2020 Ethernet BWA
 - Report - <https://bit.ly/802d3bwa2>
 - Tutorial – https://bit.ly/802d3bwa2_tut
- Reference slides in Appendix: Backup Slides

The 2020 Ethernet Bandwidth Assessment



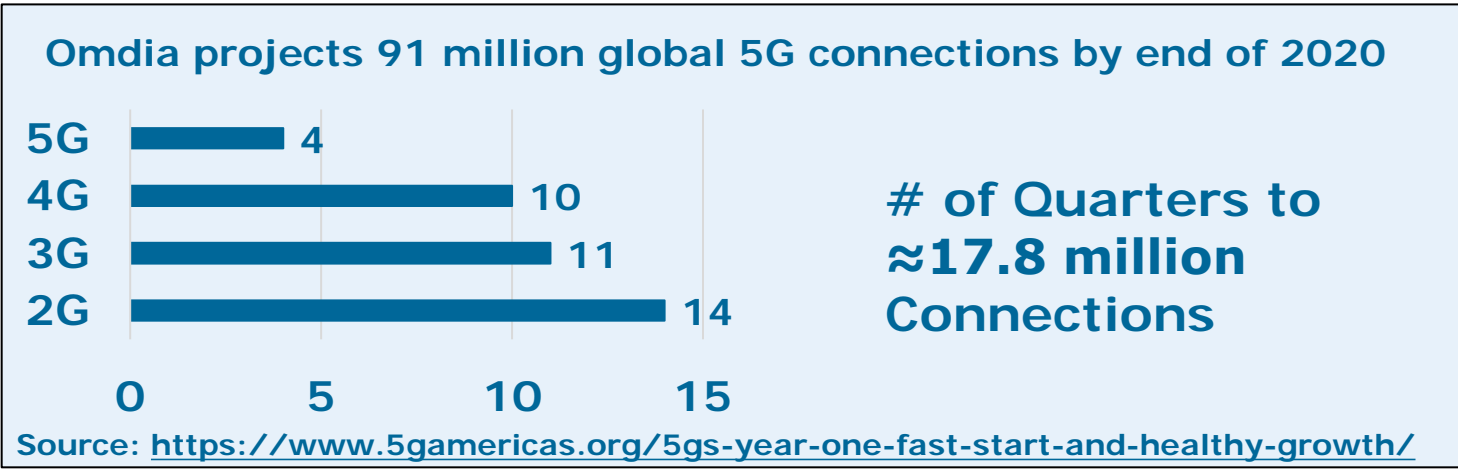
EXAMPLE EMERGING APPLICATION – 5G BACKHAUL



Source:
http://www.ieee802.org/3/B10K/public/18_01/wang_b10k_01b_0118.pdf

	LTE	LTE Advanced	5G
Africa	145	42	4
Asia & Pacific	162	74	29
Eastern Europe	93	59	14
Latin America & Caribbean	127	50	8
Middle East	44	29	12
U S & Canada	20	11	7
Western Europe	88	70	31
Global Totals	683	335	105

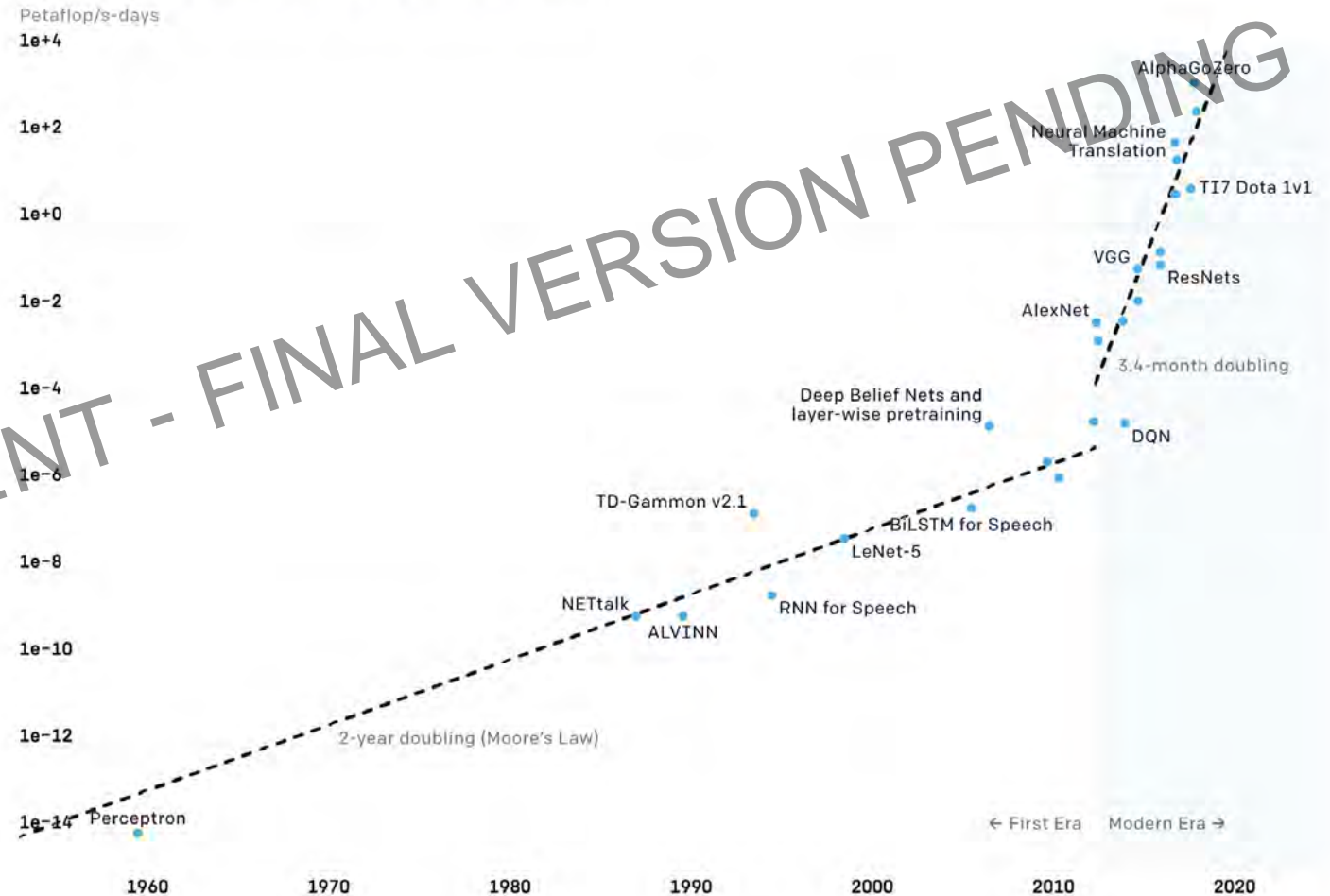
Source: as of 8/14/2020, <https://www.5gamericas.org/resources/deployments/>



ARTIFICIAL INTELLIGENCE & COMPUTE

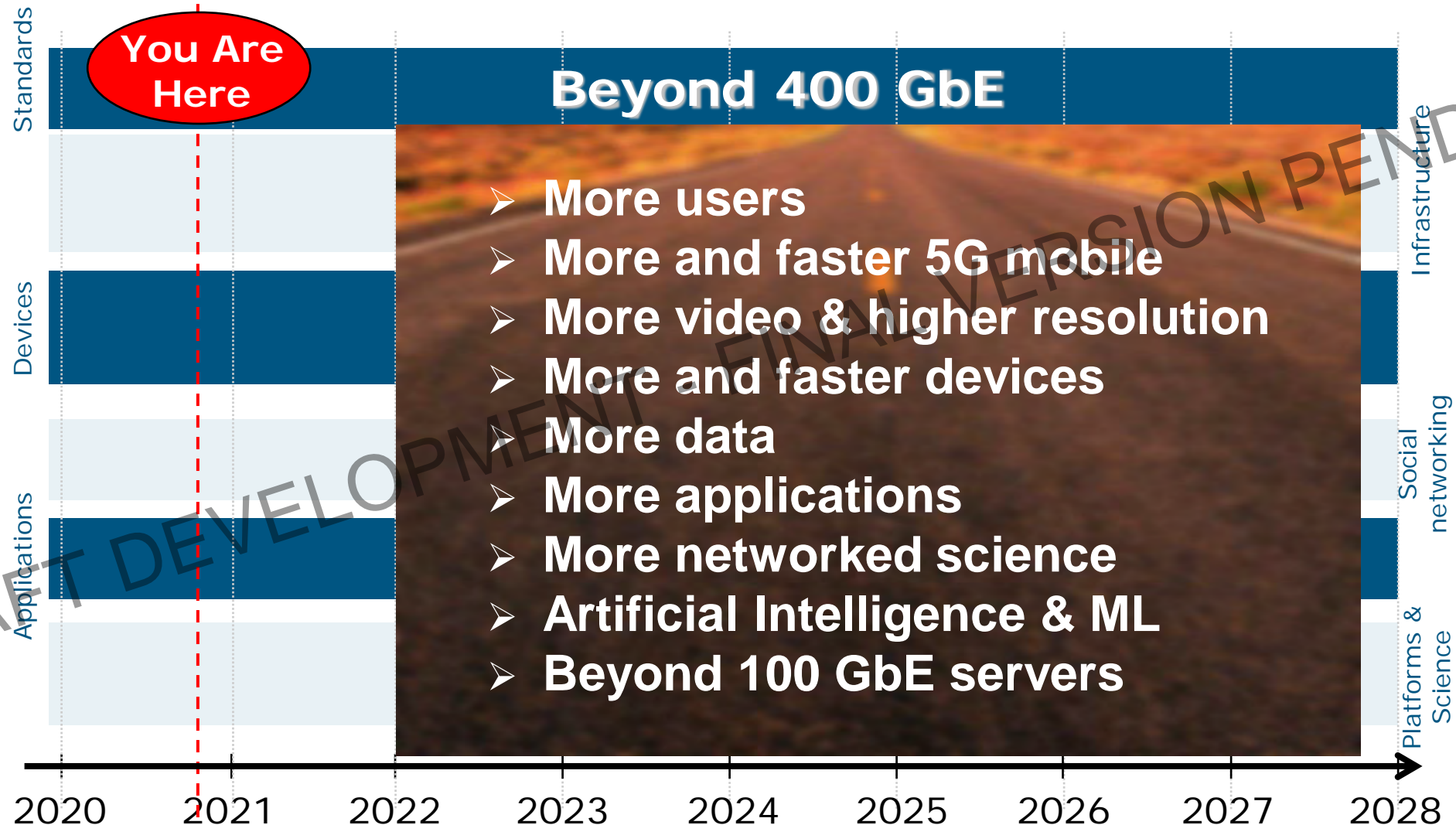
- **First Era (Before 2012)**
 - Moore's Law – 2-year doubling
 - Uncommon to use GPUs for machine learning
- **Modern Era (2012 and later)**
 - 2012 – 2014: most results used 1-8 GPUs rated at 1-2 TFLOPS
 - 2014 – 2016: large-scale results used 10-100 GPUs rated at 5-10 TFLOPS
 - 2016 – 2017: greater algorithmic parallelism (huge batch sizes, architecture search, expert iteration), specialized hardware (TPUs), faster interconnects

Two Distinct Eras of Compute Usage in Training AI Systems

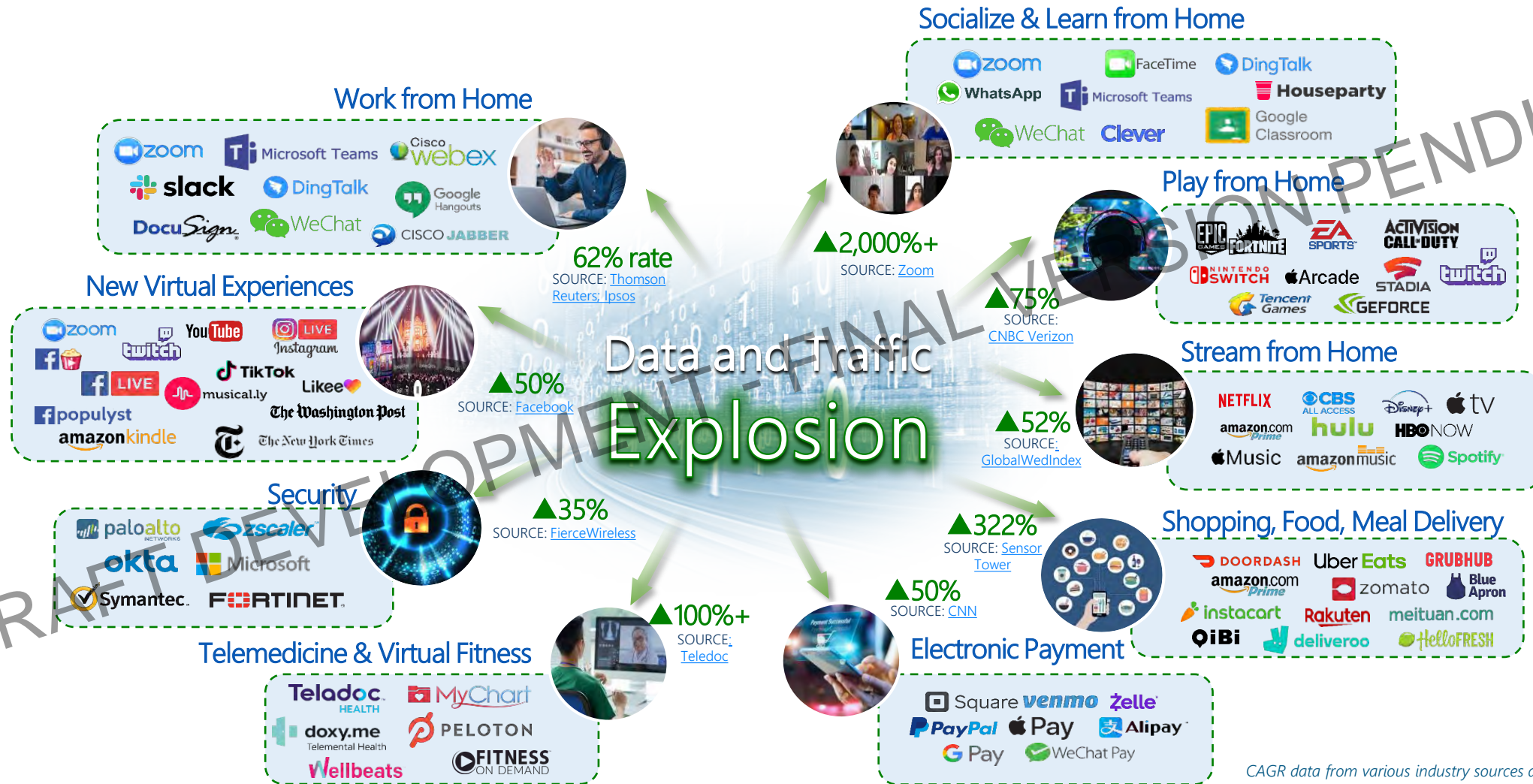


Source – OpenAI blog post 'AI and Compute' addendum 'Compute used in older headline results' posted 7th November 2019 by Girish Sastry, Jack Clark, Greg Brockman and Ilya Sutskever <<https://openai.com/blog/ai-and-compute/>>.

MORE OF THE SAME.....



COVID-19 TRENDS, APRIL 2020



CAGR data from various industry sources and Inphi estimates

Source - Inphi blog post 'Bandwidth in the Age of COVID-19' posted 22nd April 2020 by Ford Tamer, President and CEO, Inphi Corporation <<https://www.inphi.com/blog/>>

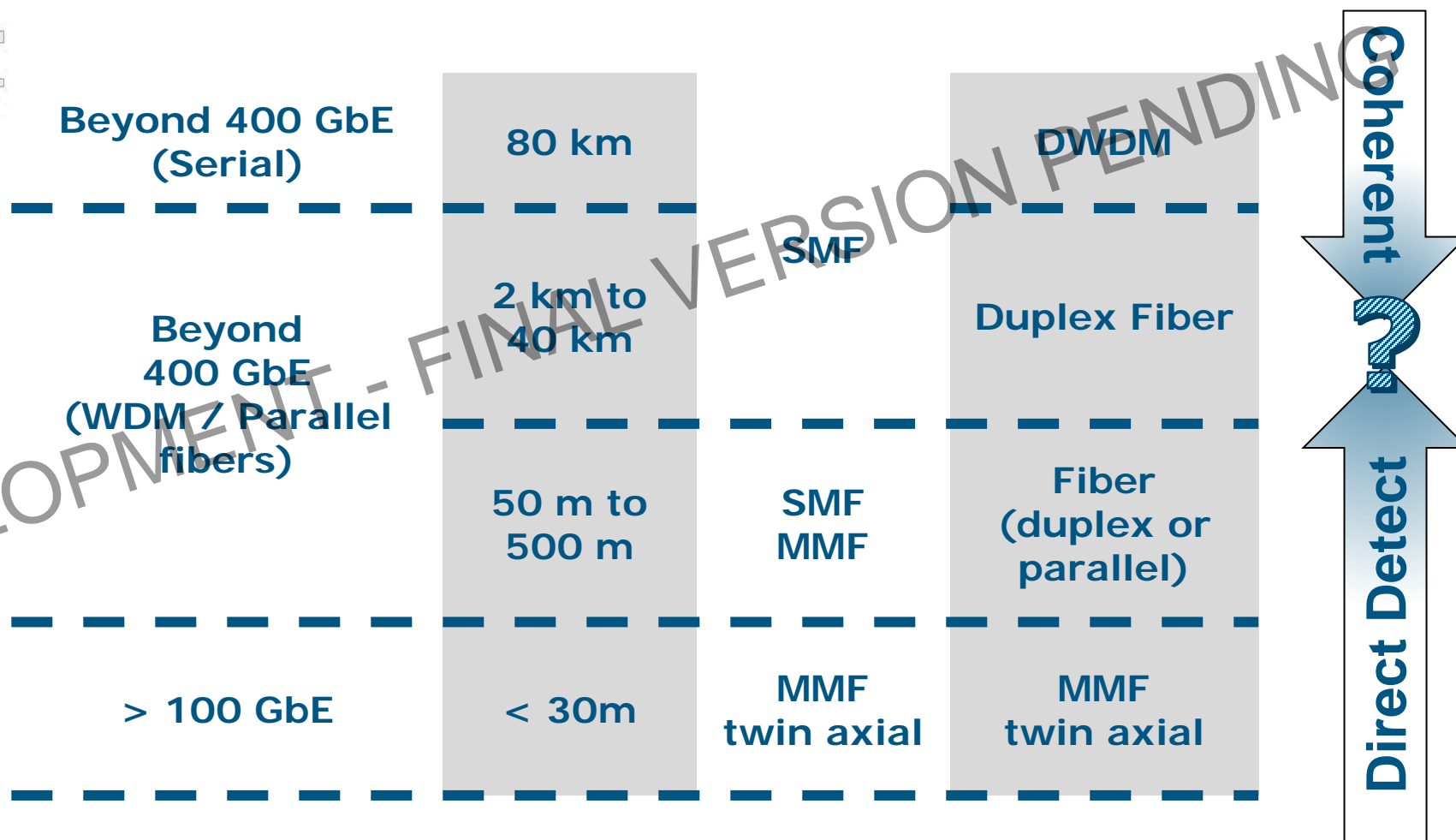
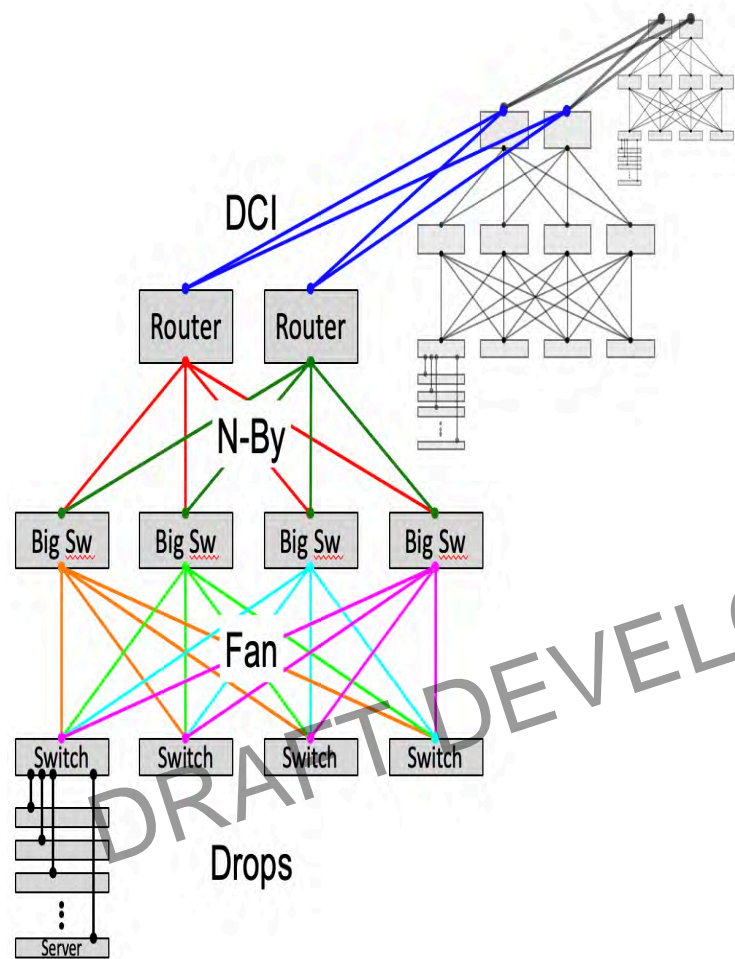
SUMMARY

- **Bandwidth growth continues and underlying factors indicate further bandwidth growth**
 - **Video and mobile!**
 - **Increasing delta between “peak” and “average”**
- **New applications fueling bandwidth growth**
- **In today's COVID-19 world**
 - **Connectivity has been critical!**
 - **“Instantaneous” growth in multiple application spaces**
- **“Up and to the right” continues**

THE TECHNICAL ROADMAP TO BEYOND 400 GbE



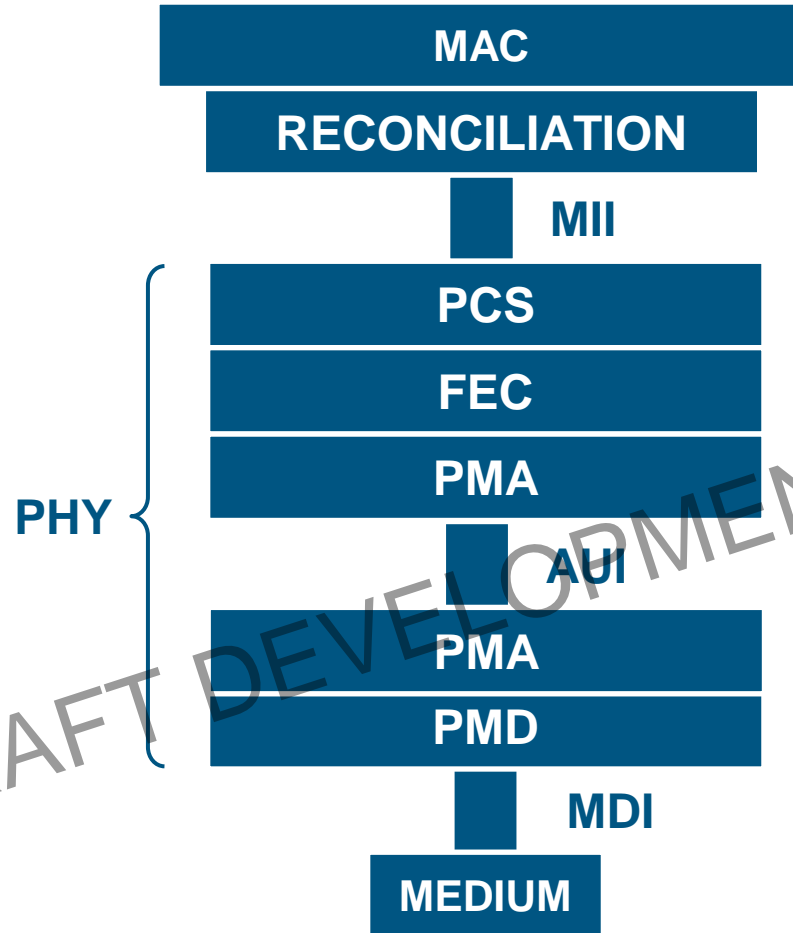
Understanding the Typical Physical Challenges



Beyond 400 GbE C2C / C2M AUI Development

Figure courtesy Jim Theodoras, HG Genuine

THE CHALLENGES TO BEYOND 400 GBE



Forward Error Correction

- AU1 Specific?
- PHY Specific?
- Multiple variants? Architecture impact?

Electrical Functions

- Increase interface channel count?
- Increase interface rate?
- Increase interface modulation order?

Optical Functions

- Increase interface channel count?
- Increase interface rate?
- Increase interface modulation order?

Media

- Increase interface channel count?
- Increase fiber count?
- Increase lambda count?
- Breakout?

MAC/PCS Technical Feasibility



- The options below are very feasible in near term technology (as an example, actual rate(s) are TBD)

MAC Rate	Technology Node	Device Type	Bus Width	Clock Rate
800 Gb/s	5 nm	ASIC	1024b	800 MHz
	5 nm	ASIC	512b	1.6 GHz
	7 nm	FPGA	1536b	533 MHz
1.6 Tb/s	5 nm	ASIC	2048b	800 MHz
	5 nm	ASIC	1024b	1.6 GHz
	5 nm (or equiv)	FPGA	3072b	533 MHz

Source – Mark Gustlin, Cisco

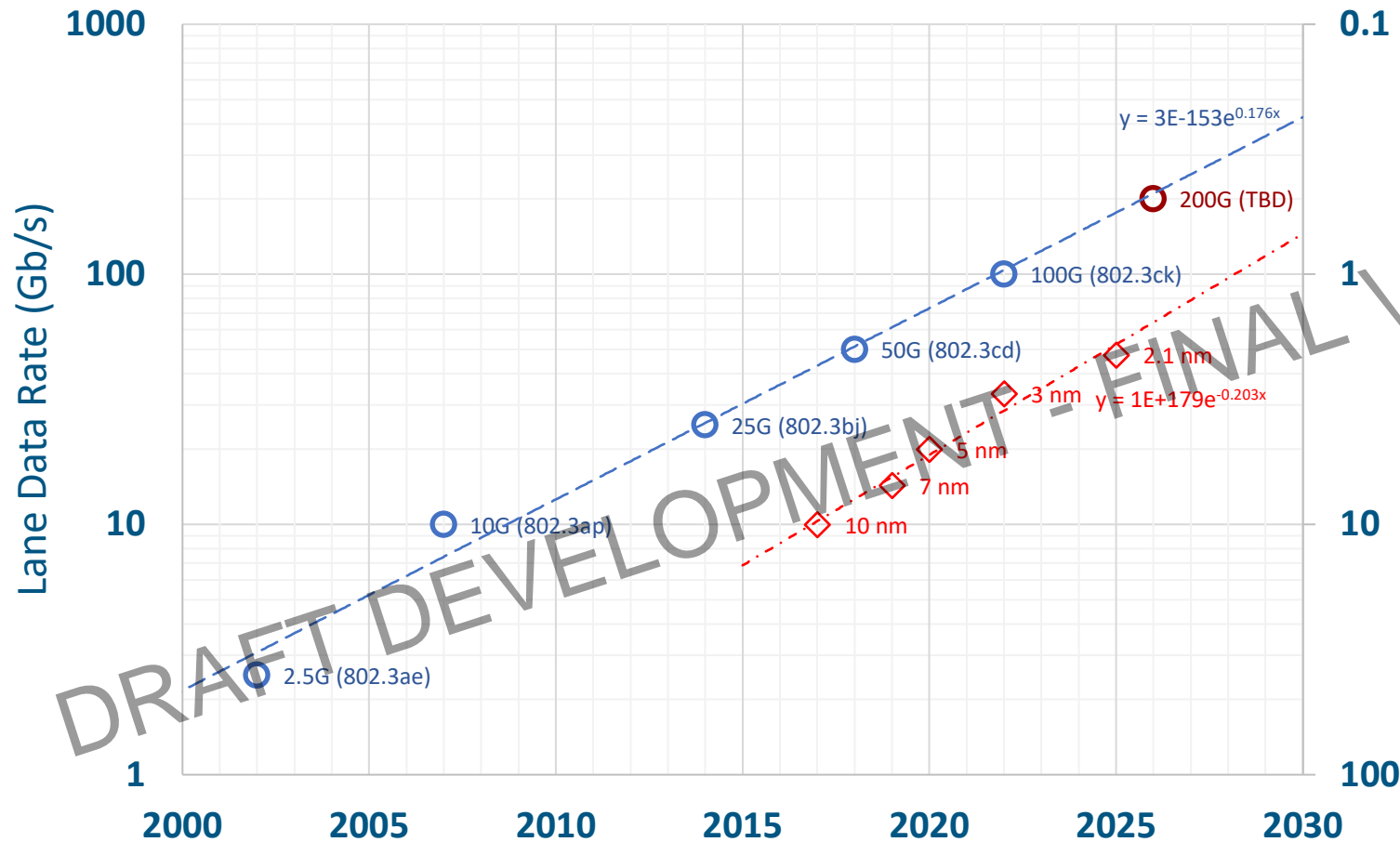
PCS/FEC

- Previous PCS concepts could be re-used
 - 64b/66b, transcoding, scrambling, AMs
- Will likely want a new stronger FEC for 200 Gb/s lane (if the project chooses to define 200 Gb/s per lane)
 - Multiple FEC options for direct detect, coherent light and longer reach coherent?
 - Still support end to end FEC for some options?
 - Optimize gain, latency, power and implementation burden for chosen FECs
 - While minimizing the overall number of FEC options

Source – Mark Gustlin, Cisco

CMOS Roadmap

Comparison of Lane Data Rate and Node Label Timelines



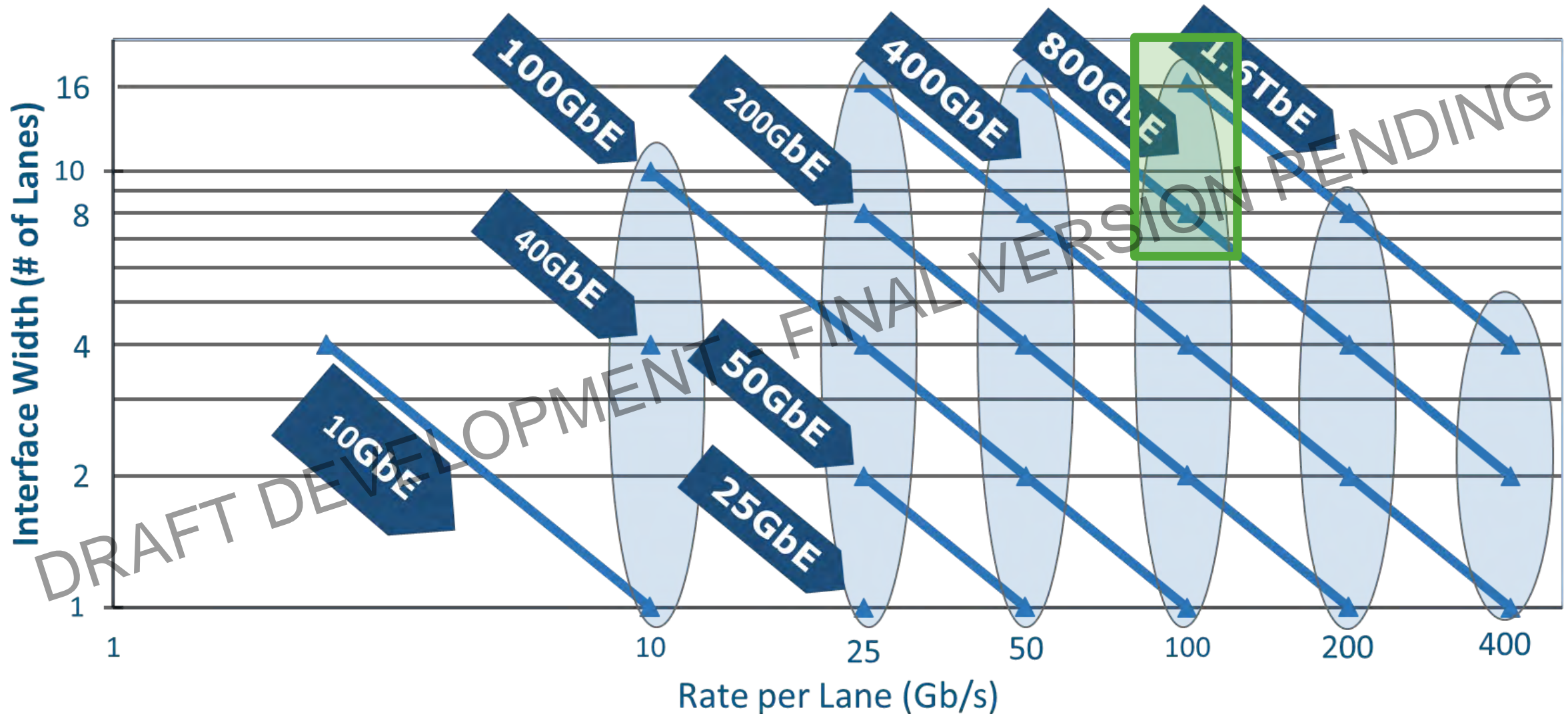
- The upper data (blue) shows evolution of electrical lane data rate over time.
- The lower data (red) shows the evolution of node label over time.
- Current designs for 100 Gb/s per lane are in 7 nm and are moving to 5 nm.
- 3 nm and 2.1 nm will be available when 200 Gb/s per lane is standardized.
- The node label (halving every 3.4 years) is progressing faster than the electrical lane rate (doubling every 3.9 years).

Source – Matt Brown, Huawei Canada

DSP Architecture Advances

DRAFT DEVELOPMENT - FINAL VERSION PENDING

Beyond 400 GbE - Leveraging 100 Gb/s



Industry Efforts - 100 Gb/s Signaling

➤ IEEE 802.3

➤ Standard – IEEE P802.3bs – 400GBASE-DR4 (4x100G)

➤ In Development

- IEEE P802.3ck 100 Gb/s, 200 Gb/s, and 400 Gb/s Electrical Interfaces Task Force
- IEEE P802.3cu 100 Gb/s and 400 Gb/s over SMF at 100 Gb/s per Wavelength Task Force
- IEEE P802.3db 100 Gb/s, 200 Gb/s, and 400 Gb/s Short Reach Fiber Task Force

➤ Other Industry Efforts

- OIF Common Electrical Interface 112G Efforts
- 100G Lambda MSA (100Gb/s optical interfaces specifications)

800 Gb/s Industry Activities

➤ Ethernet Technology Consortium

- <https://ethernettechnologyconsortium.org/>
- "The 800 GbE specification introduces a new media access control (MAC) and Physical Coding Sublayer (PCS)"

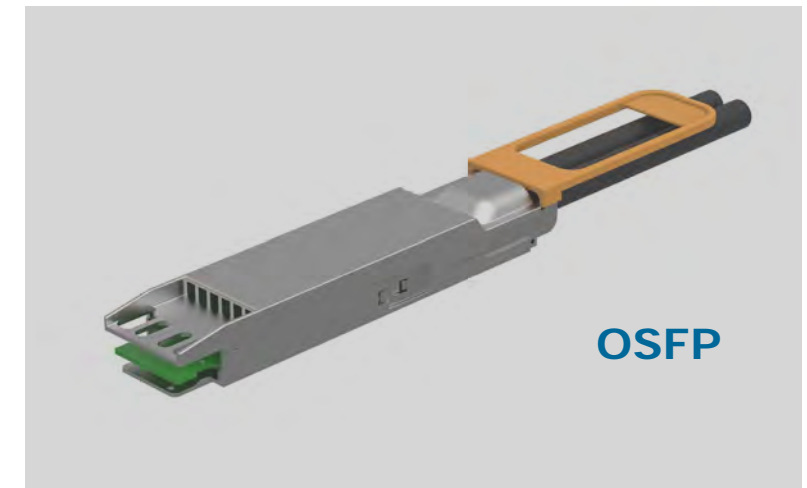
➤ QSFP-DD800 MSA

- <http://www.qsfp-dd800.net/>
- Rev 1.0 released Mar 6 2020

➤ OSFP

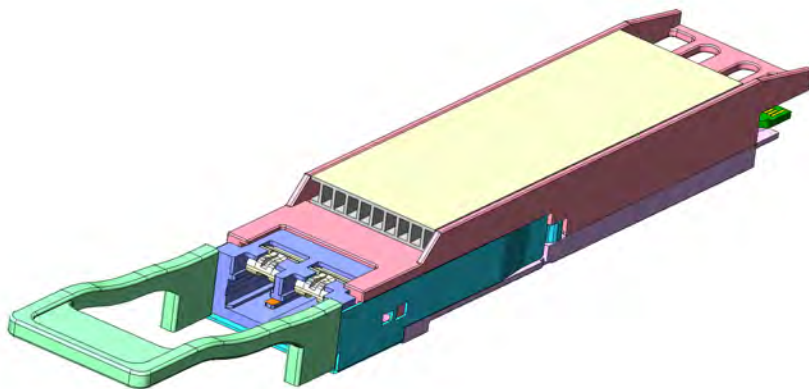
➤ 800G Pluggable MSA

- <https://www.800gmsa.com/>
- 800G PSM8 specification (Draft 1.0) - Specification covering cost effective 8x100G transmission over at least 100m



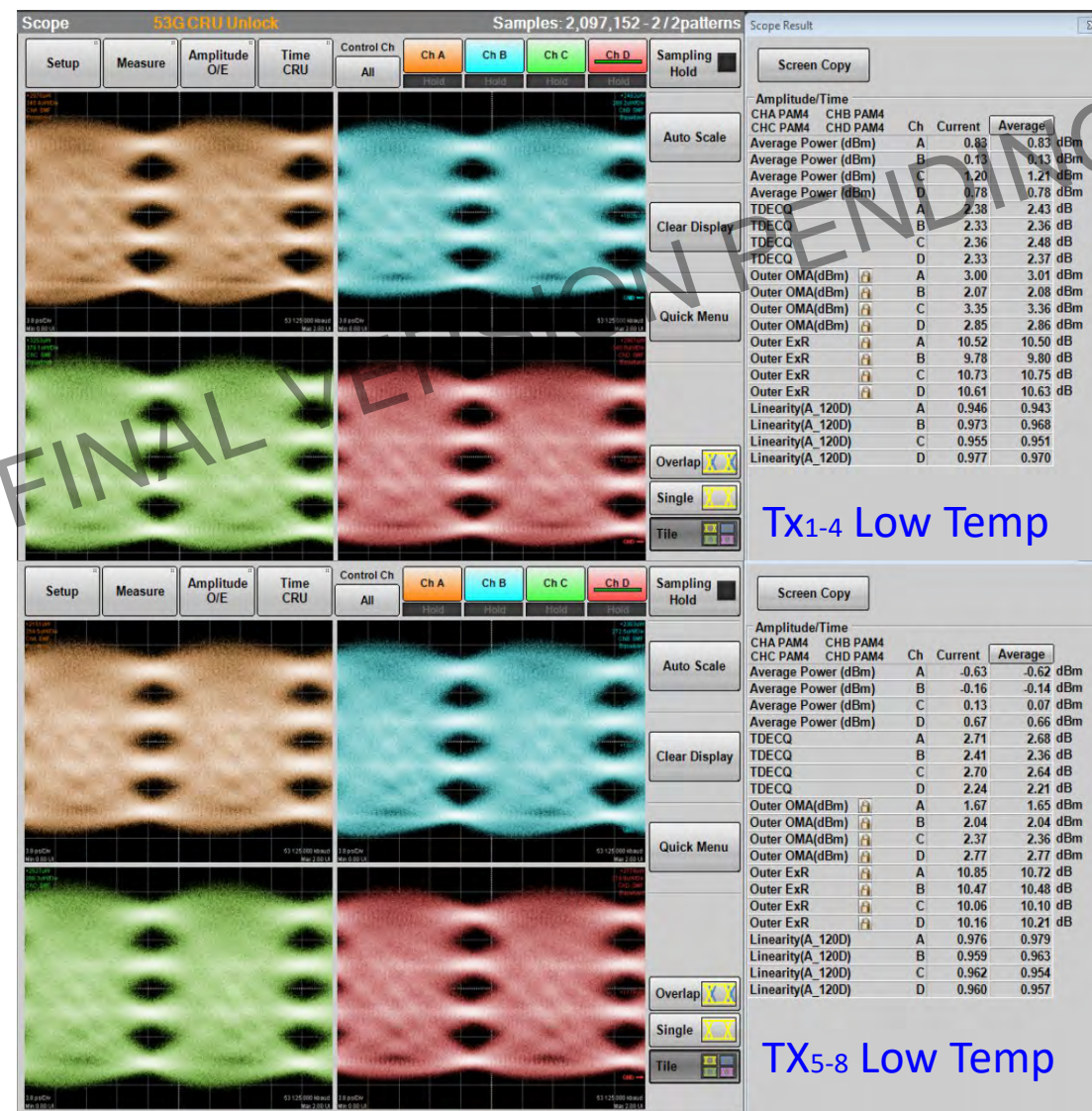
Source— Nathan Tracy, TE Connectivity

Example: 800 Gb/s OSFP Capacity Module

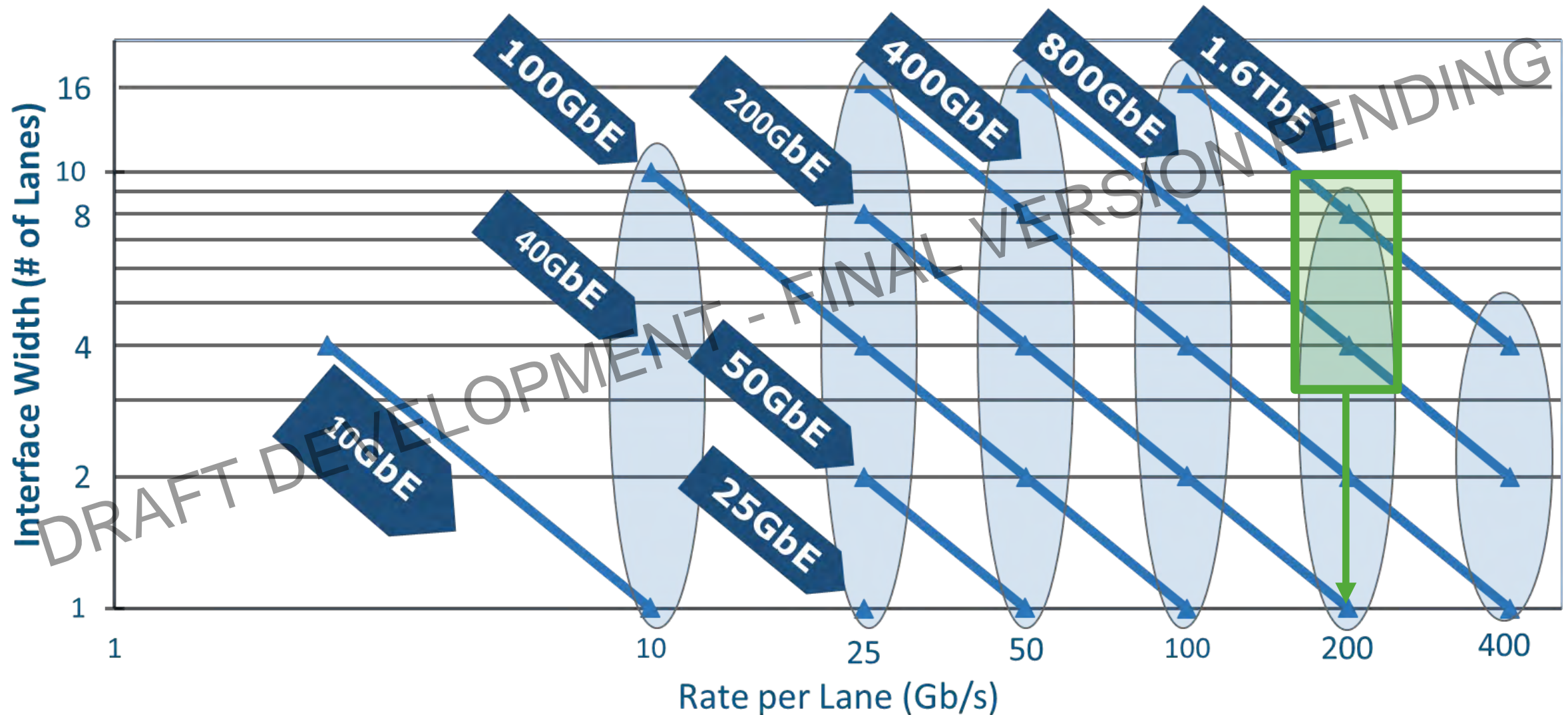


- OSFP Form Factor
- 8x100G DR8+ 2km with MPO-16 and 2x400G FR4 with QS connector
- OIF CEI-112G-VSR interface
- PMD spec follows 400G DR4+ and FR4. interoperable with 400G
- 0~70degC 18W, 10~60C 17W
- 7nm DSP inside

Source – Tedros Tsegaye, Innolight



Beyond 400 GbE - Leveraging 200 Gb/s



Beyond 100 Gb/s Research is Underway

- S. Yamaoka et al., "239.3-Gbit/s net rate PAM-4 transmission using directly modulated membrane lasers on high-thermal-conductivity SiC" in Proceedings of European Conference on Optical Communication (ECOC), 2019/9.
- X. Pang et al., 200 Gbps/lane IM/DD Technologies for Short Reach Optical Interconnects, <https://core.ac.uk/download/pdf/289286726.pdf>, 2019/04/24.
- W. Heni et al., Ultra-High-Speed 2:1 Digital Selector and Plasmonic Modulator IM/DD Transmitter Operating at 222 GBaud for Intra-Datacenter Applications, <https://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-38-9-2734>, 2020/9.
- S Lange et al., 100 GBd Intensity Modulation and Direct Detection with an InP-based Monolithic DFB Laser Mach-Zehnder Modulator, Journal of Lightwave Technology, https://www.researchgate.net/publication/319259046_100_GBd_Intensity_Modulation_and_Direct_Detection_with_an_InP-based_Monolithic_DFB_Laser_Mach-Zehnder_Modulator, 2017/8.
- E. Sentieri et al., "12.2 A 4-Channel 200Gb/s PAM-4 BiCMOS Transceiver with Silicon Photonics Front-Ends for Gigabit Ethernet Applications," 2020 IEEE International Solid- State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2020, pp. 210-212, doi: 10.1109/ISSCC19947.2020.9062992.
- T. Wettlin et al., "Beyond 200 Gb/s PAM4 transmission using Tomlinson-Harashima precoding," 45th European Conference on Optical Communication (ECOC 2019), Dublin, Ireland, 2019, pp. 1-4, doi: 10.1049/cp.2019.0834.
- Net 212.5 Gbit/s Transmission in O-band With a SiP MZM, One Driver and Linear Equalization, Maxime Jacques¹, Zhenping Xing¹, Alireza Samani¹, Xueyang Li¹, Eslam El-Fiky¹, Samiul Alam¹, Olivier Carpentier¹, Ping-Chiek Koh², David Plant¹; ¹McGill Univ., Canada; ²Lumentum, USA. OFC-2020, Post deadline paper Th4A.3

200 Gb/s Signaling – The Next Generation?

- OIF approves CEI 224G Development Project

- <https://www.businesswire.com/news/home/20200826005437/en/OIF-Approves-CEI-224G-Development-Project-Reviews-Co-packaging>

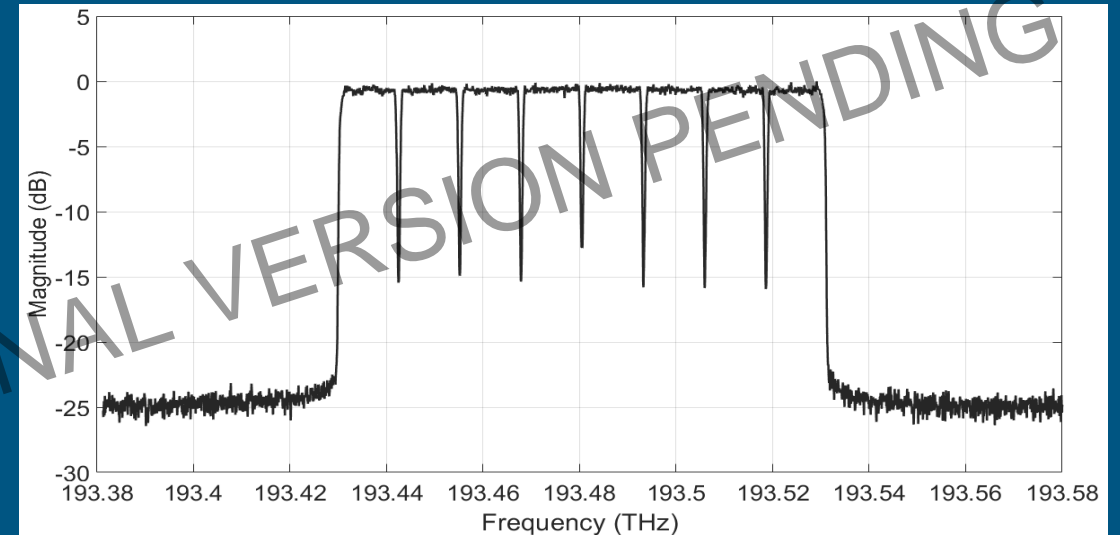
DRAFT DEVELOPMENT - FINAL VERSION PENDING

800 Gb/s Single Wavelength Transmission

The Future of Coherent is emerging

- Successful trial of 800 Gb/s single-wave transmission over 950 km - <https://bit.ly/2Wdkh8e>
- Platform supporting 200 Gb/s to 800 Gb/s single-carrier - <https://bit.ly/2KLpW05>
- "Industry's first 800G tunable ultra-high-speed optical module" <https://bit.ly/2yTYNFK>
- "Verizon says it has successfully transmitted an 800-Gb/s wavelength on its live network" - <https://bit.ly/3d2GX1M>

800 Gb/s single wavelength transmission over 730km in real world long-haul network



8 subcarrier constellation

<https://www.lightreading.com/optical-ip/infinera-windstream-tout-optical-networking-milestone/d/d-id/761738>

Source – Ted Sprague, Infinera

SUMMARY

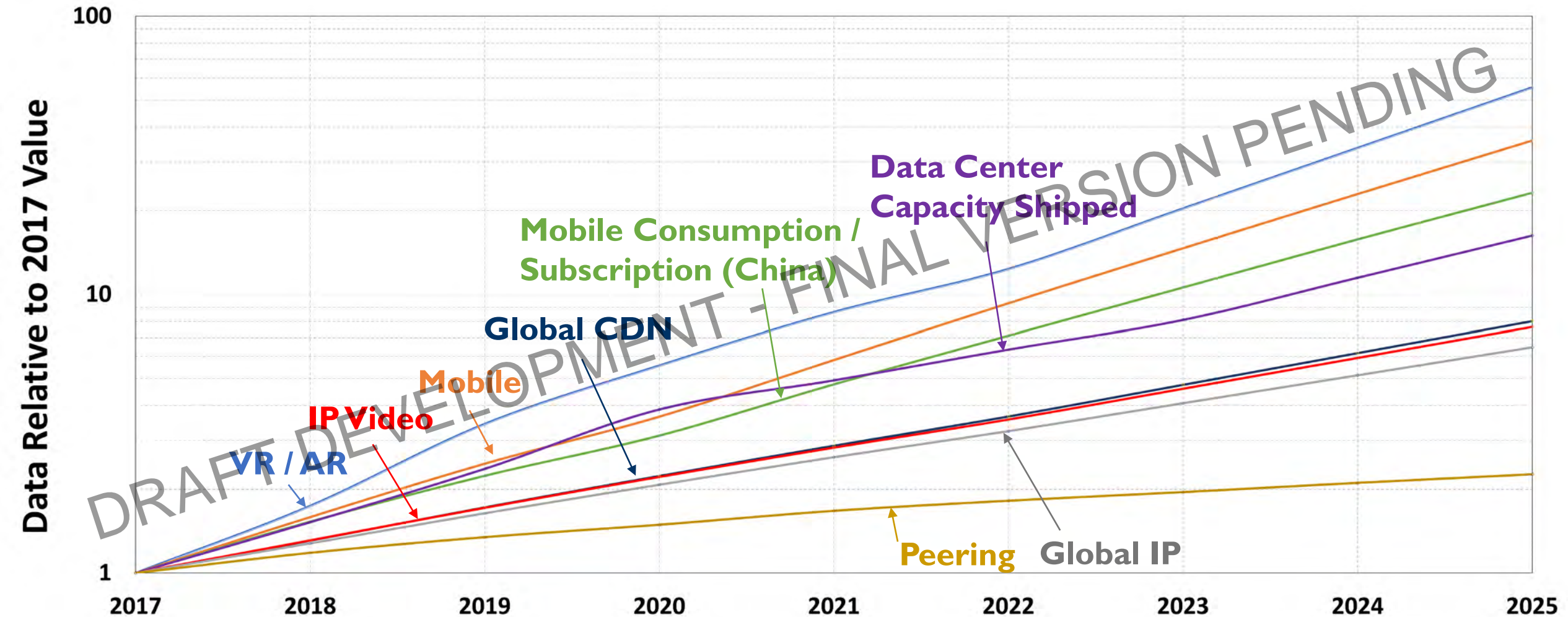
- Path to Beyond 400 GbE exists
- Leverage 100 Gb/s building blocks
- 800 GbE building blocks and example available now
- Plausible implementations for today and next generation
- 800 Gb/s over a single wavelength for DWDM systems is emerging now

BEYOND 400 GbE WHY NOW?

DRAFT DEVELOPMENT - FINAL VERSION PENDING

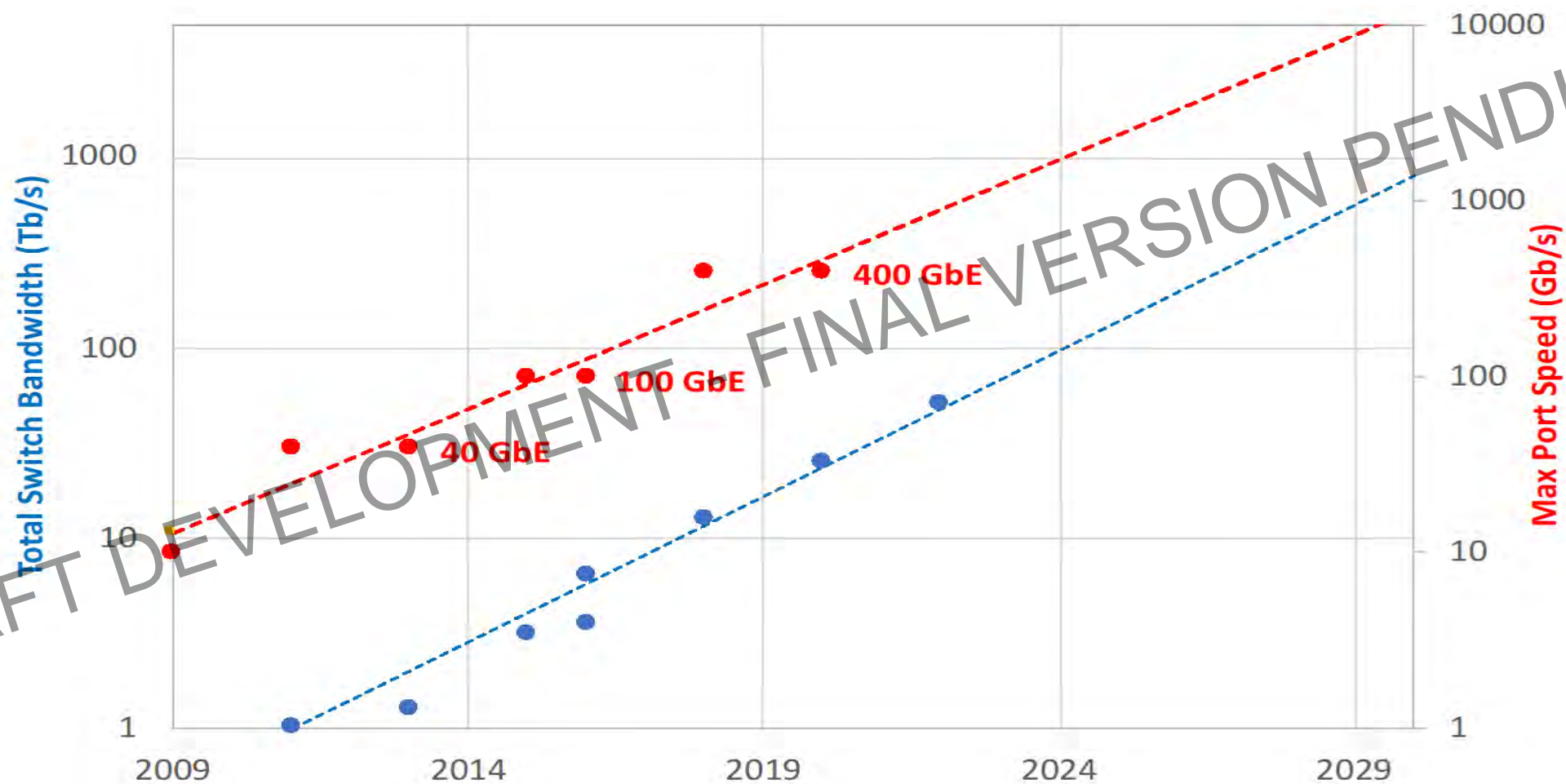


The 2020 Ethernet Bandwidth Assessment



Source: <https://bit.ly/802d3bwa2>

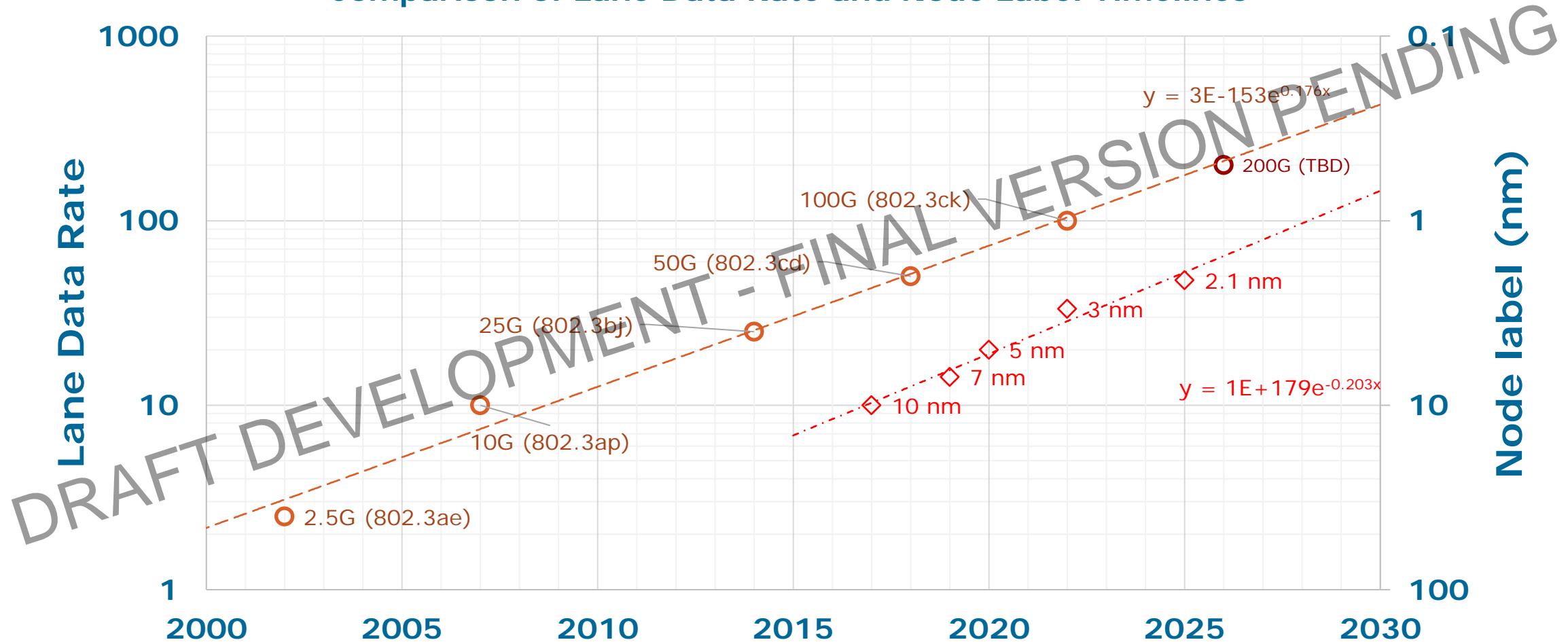
TRENDLINE – SWITCH CAPACITY



Source: Rob Stone, Facebook

TRENDLINE – SERDES DEVELOPMENT

Comparison of Lane Data Rate and Node Label Timelines



Source: Matt Brown, Huawei

Understanding the Typical Physical Challenges

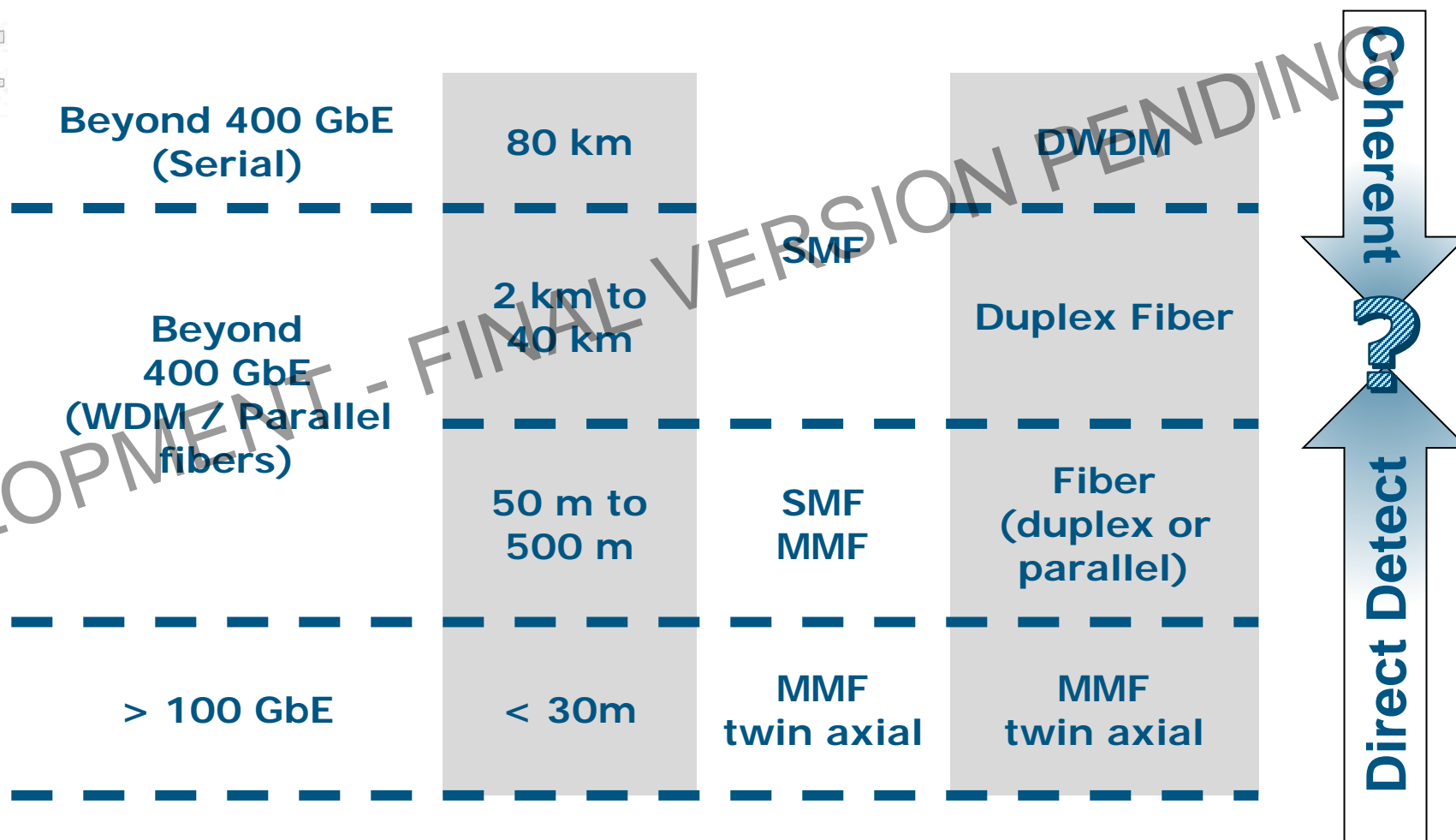
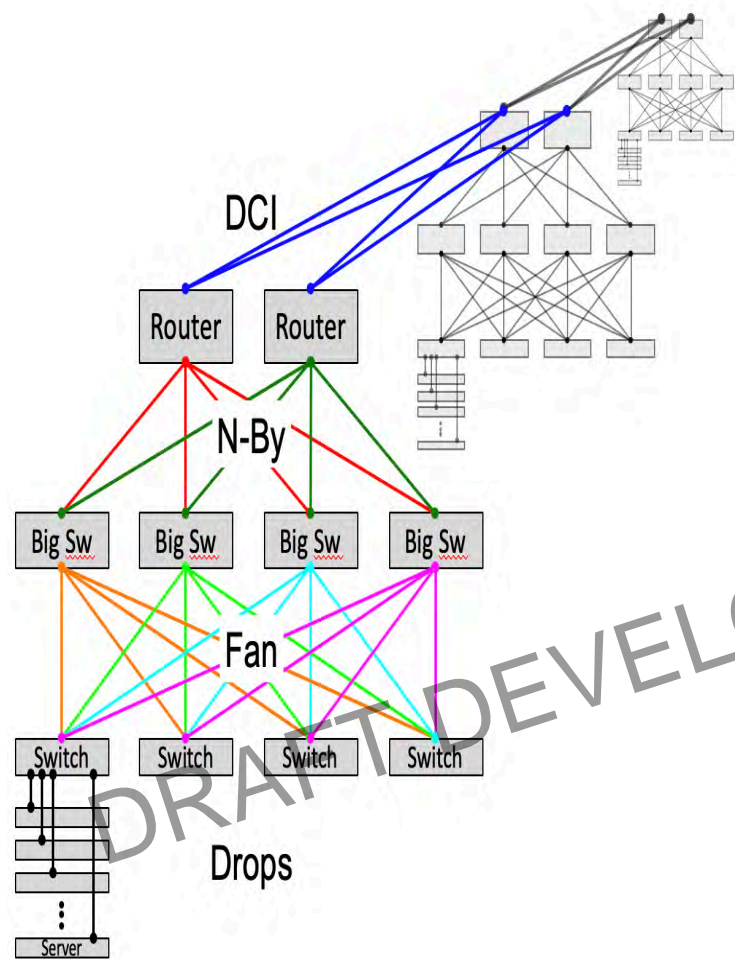
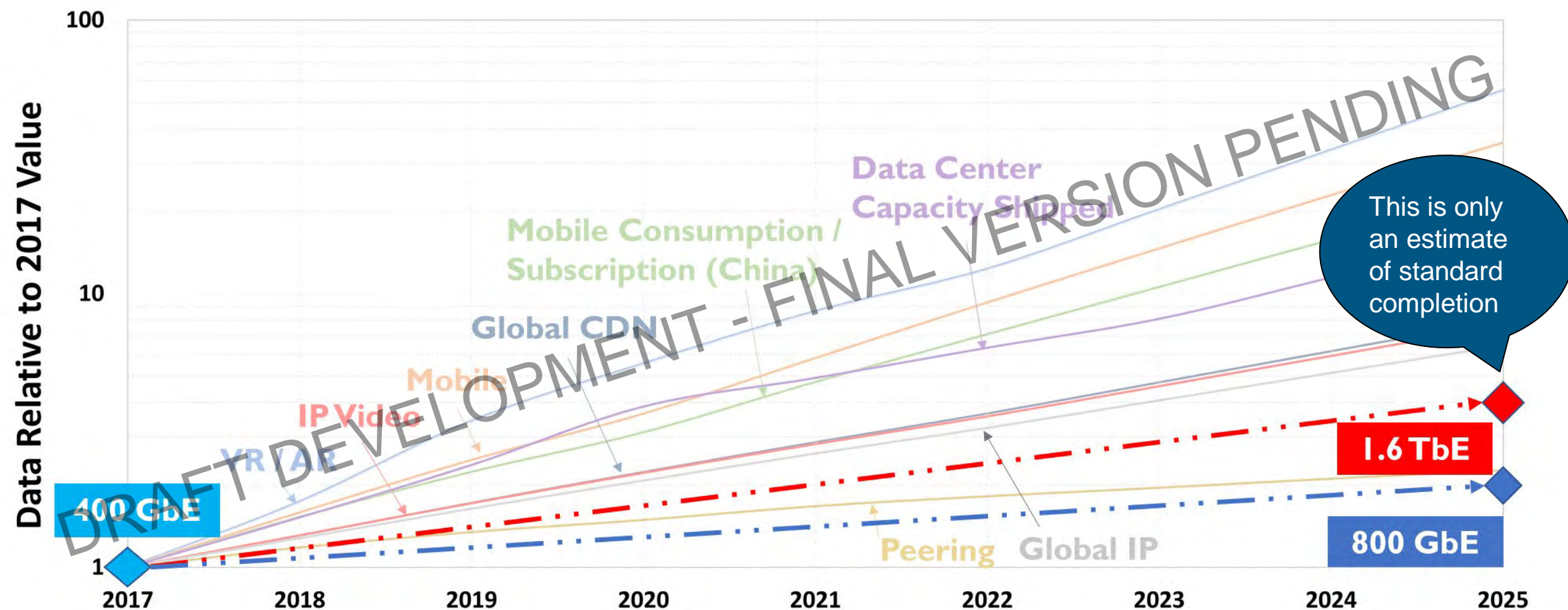


Figure courtesy Jim Theodoras, HG Genuine

CONSIDERING THE NEXT ETHERNET RATE STANDARD



Source: <https://bit.ly/802d3bwa2>

SUMMARY

- **Bandwidth –**
 - Underlying factors all indicate continued growth
 - Exponential growth continues!
- **New bandwidth generating applications constantly being introduced**
 - Mobile (5G) / Video
 - Artificial Intelligence
 - Virtual / Augmented Reality
- **Today's world stressing the need for connectivity and bandwidth**
- **Last two “Higher Speed” efforts (from CFI to standard ratification)**
 - 40 / 100 GbE – 3 years, 11 months
 - 200 / 400 GbE – 4 years, 9 months
- **There is some time between standard ratification and product introduction**
 - The bandwidth problem will only continue to grow
- **We need to begin the process to study the problem! Big questions to consider**
 - Next speed or speeds?
 - What physical layer specifications?

Proposed Study Group Chartering Motion

Approve the formation of a Beyond 400 Gb/s Ethernet Study Group to consider development of a Project Authorization Request (PAR) and Criteria for Standards Development (CSD) responses for:

- 1. Beyond 400 Gb/s Ethernet;**
- 2. Physical Layers specifications for existing Ethernet rates based on any signaling rate used for (1).**

Supporters (as of 9/14/20)

John Thananya	Abbott Baldwin	Corning Incorporated	Cedric David	Lam Lewis	Google Lumentum	Yoshiaki Massimo	Sone Sorbara	NTT GlobalFoundries
Vipul Paul	Bhatt Brooks	II-VI Incorporated	Junjie Robert	Li Lingle	China Telecom OFS	Ted Rob	Sprague Stone	Infinera Facebook
Matt Leon	Brown Bruckman	Huawei Technologies Canada	Hai-Feng Kent	Liu Lusted	HG Genuine Intel	Steve Bharat	Swanson Tailor	Corning Incorporated Semtech
Frank Ayla	Chang Chang	Source Photonics	Ilya Valerie	Lyubomirsky Maguire	Inphi Siemon	Tomoo Jim	Takahara Theodoras	Fujitsu HG Geuine USA
Weiqliang	Cheng	China Mobile	David David	Malicoat	Malicoat Networking Solutions	Nathan Tedros	Tracy Tsegaye	TE Connectivity Innolight
John	D'Ambrosia	Futurewei, U.S. Subsidiary of Huawei	Flavio Larry	Marques McMillan	Furukawa Electric Western Digital	Jeff Ed	Twombly Ulrichs	Credo Semiconductor Intel
John Claudio	DeAndrea DeSanti	II-VI Inc Dell Technologies	Rich Shimon	Mellitz Muller	Samtec	Xinyuan Yangling	Wang Wen	Huawei Futurewei
Vince Ali	Ferretti Ghiasi	Corning Incorporated	Dale Shawn	Murray Nicholl	LightCounting Xilinx	Zhao Chongjin	Wenyu Xie	CAICT Alibaba
Joel Bob	Goergen Grow	Cisco RMG Consulting	Paul Tom	Nikolich Palkert	Independent Samtec	Shuto James	Yamamoto Young	NTT Commscope
Mark Rubio	Gustlin Han	Cisco China Mobile	Earl Jerry	Parsons Pepper	CommScope Keysight Technologies	Xu Xiang	Yu Zhou	Huawei Google
Xiang Tom	He Issenhuth	Huawei	David Rick	Piebler Pimpinella	Dell Technologies Panduit	Yan George	Zhuang Zimmerman	Huawei CME Consulting
Ken	Jackson	Sumitomo Electric Device Innovations USA	Rick Ed	Rabinovich Sayre	Keysight Technologies NESA	Pavel	Zivny	Tektronix
John Lokesh	Johnson Kabra	Broadcom Synopsys	Priyank	Shukla	Synopsys			
Mark	Kimber	Semtech						

STRAW POLLS

DRAFT DEVELOPMENT - FINAL VERSION PENDING



Call-for-interest

- **Should a Study Group be formed for “Beyond 400 Gb/s Ethernet”**

- **YES**

- **No**

- **Abstain**

- **Call Count**

participation

- I would participate in the “Beyond 400 Gb/s Ethernet” Study Group in IEEE 802.3
 - Tally:
- I believe my affiliation would support participation in the “Beyond 400 Gb/s Ethernet” Study Group in IEEE 802.3
 - Tally:

Future work

- Ask 802.3 WG for approval
- If approved, request formation of “Beyond 400 Gb/s Ethernet” Study Group by 802 EC
- If approved,
 - Creation of Study Group page /reflector
 - First Study Group meeting [teleconference?] anticipated for Jan 21 Interim

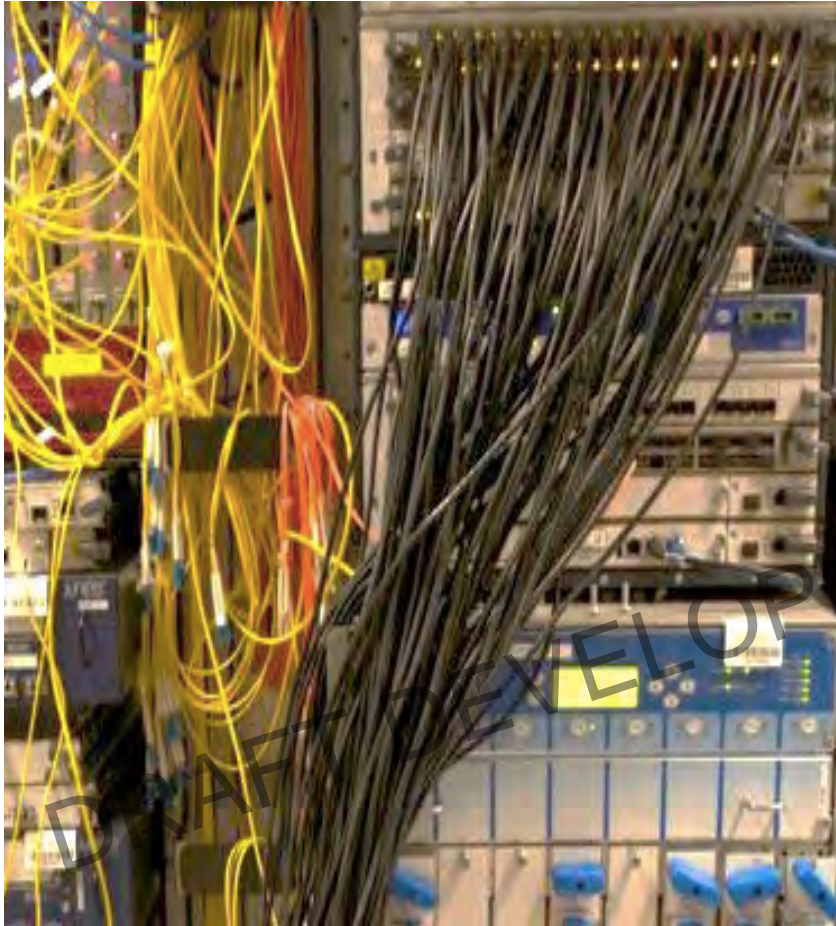


APPENDIX: BACKUP SLIDES

DRAFT DEVELOPMENT - FINAL VERSION PENDING



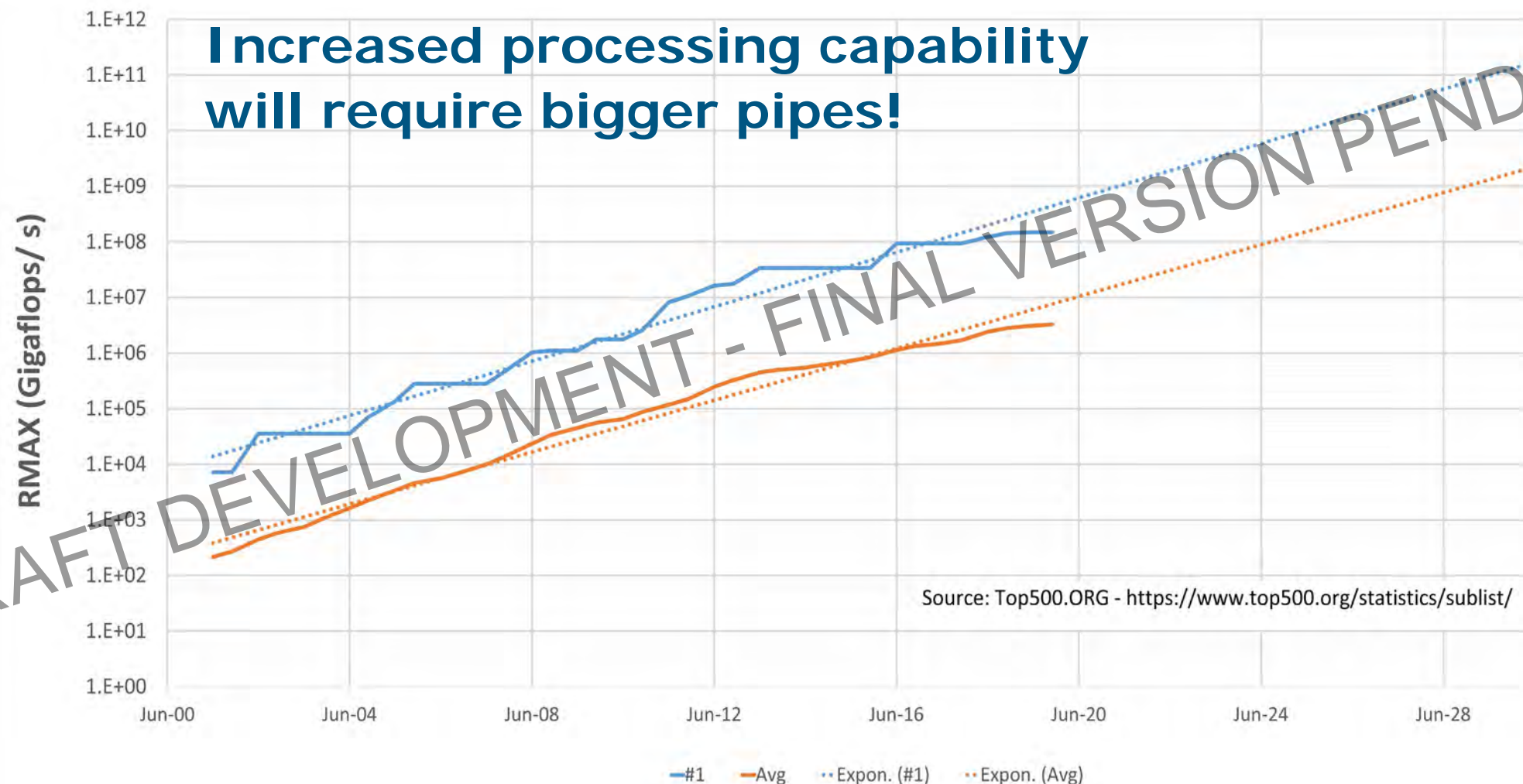
LINK AGGREGATION WILL NOT SUFFICE



Courtesy, David Ofelt, Juniper.

- Problem: Need to scale the Network (density & cost)
- Temporary Solution: Link Aggregation
- Pros:
 - Addresses bandwidth requirements between releases of faster links
- Cons:
 - Non-deterministic performance
 - Fastest flow limited to individual link speed
 - Exponential bandwidth growth implies:
 - Exponential growth in number of links
 - Growth in operational & management issues
 - Doesn't scale forever.
- **Faster links address these issues and they will be LAGGed!**

HIGH PERFORMANCE COMPUTING



WORLD INTERNET USAGE

Total World	As of 3/31/19 ¹	As of 12/31/19 ³	Increase	As of 7/20/20 ²	Increase
Population	7,716,223,209	7,796,615,710	80,392,501	7,796,949,710	80,726,501
Internet Users	4,383,810,342	4,574,150,134	190,339,792	4,833,521,806	449,711,464
Internet Penetration	57%	59%	2%	62%	5%

Top 20 Countries	As of 3/31/19 ¹	As of 12/31/19 ³	Increase
Population	5,187,499,066	5,233,377,837	45,878,771
Internet Users	3,117,533,898	3,241,273,512	123,739,614
Internet Penetration	60%	62%	2%

Rest of World	As of 3/31/19 ¹	As of 12/31/19 ³	Increase
Population	2,565,984,143	2,563,237,873	-2,746,270
Internet Users	1,229,027,955	1,332,876,622	103,848,667
Internet Penetration	48%	52%	4%

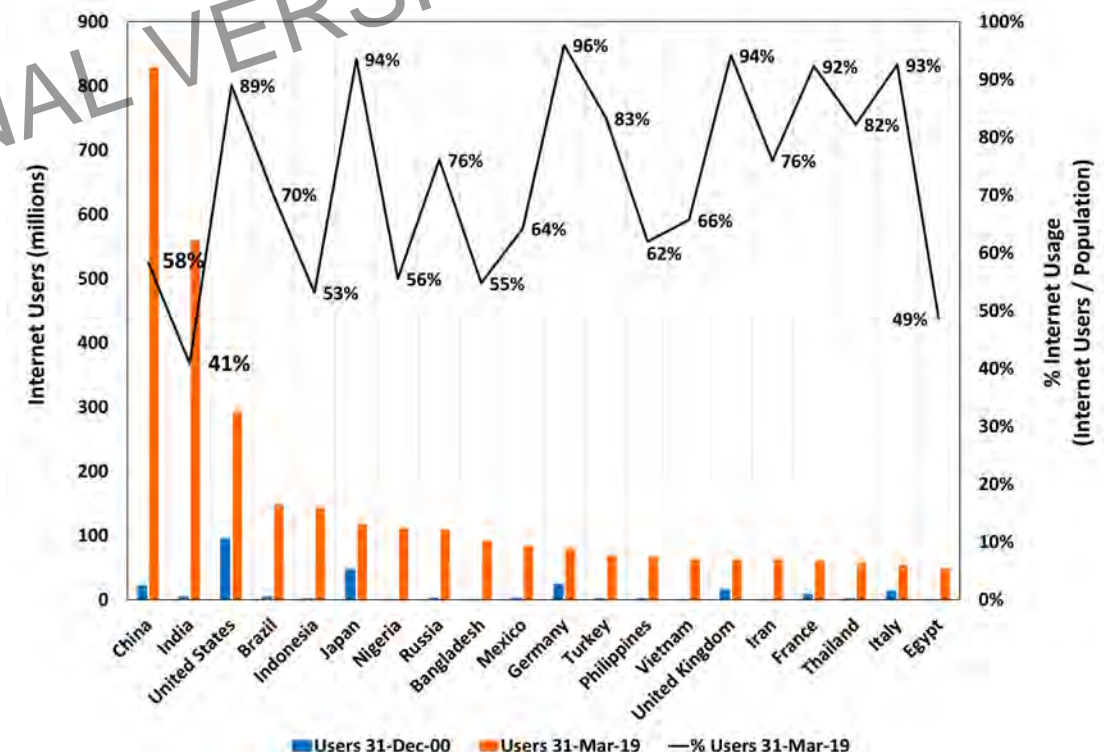
Observations

- ❖ Only 8 countries had at least 80% connectivity
- ❖ ≈ 450 million users increase
- ❖ 5% increase in Total World Internet Penetration since Mar 31 2019

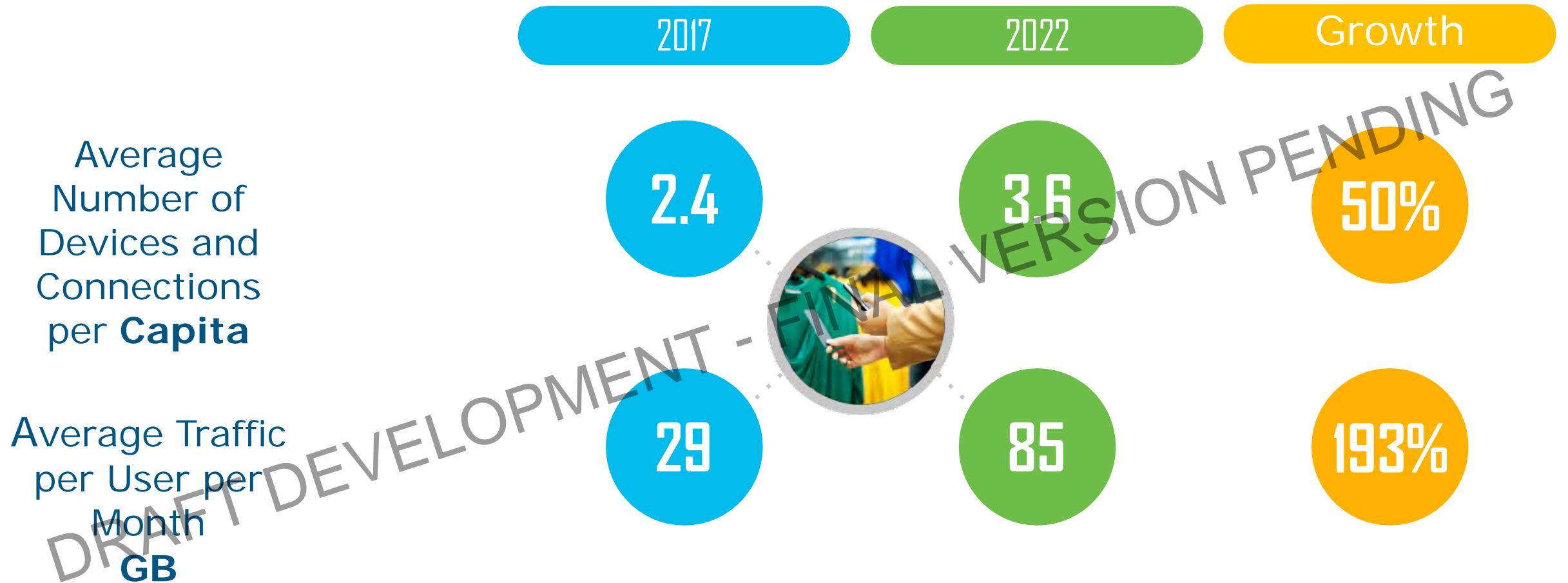
1. IEEE 802.3 BWA, PART II

2. [HTTPS://WWW.INTERNETWORLDSTATS.COM/STATS.HTM](https://www.internetworldstats.com/stats.htm)

3. [HTTPS://WWW.INTERNETWORLDSTATS.COM/STATS.HTM](https://www.internetworldstats.com/top20.htm)



GLOBAL DEVICES / CONNECTIONS AVERAGE PER CAPITA



Number of connected devices per capita is growing
The average traffic per user is growing at a much faster rate

Source: Cisco VNI Forecast Update, http://www.ieee802.org/3/ad_hoc/bwa2/public/calls/19_0624/nowell_bwa_01_190624.pdf

GLOBAL DEVICE CONNECTION GROWTH (AVERAGE)

North America			
(Mb/s)	2017	2022	CAGR
Fixed Broadband	43.2	94.2	16.9%
Wi-Fi	37.1	83.8	17.7%
Cellular	16.3	42.0	20.8%

Western Europe			
(Mb/s)	2017	2022	CAGR
Fixed Broadband	37.9	76.0	14.9%
Wi-Fi	25.0	49.5	14.6%
Cellular	16.0	50.5	25.8%

Central & Eastern Europe			
(Mb/s)	2017	2022	CAGR
Fixed Broadband	32.8	46.7	7.3%
Wi-Fi	19.5	32.8	11.0%
Cellular	10.1	26.2	21.0%

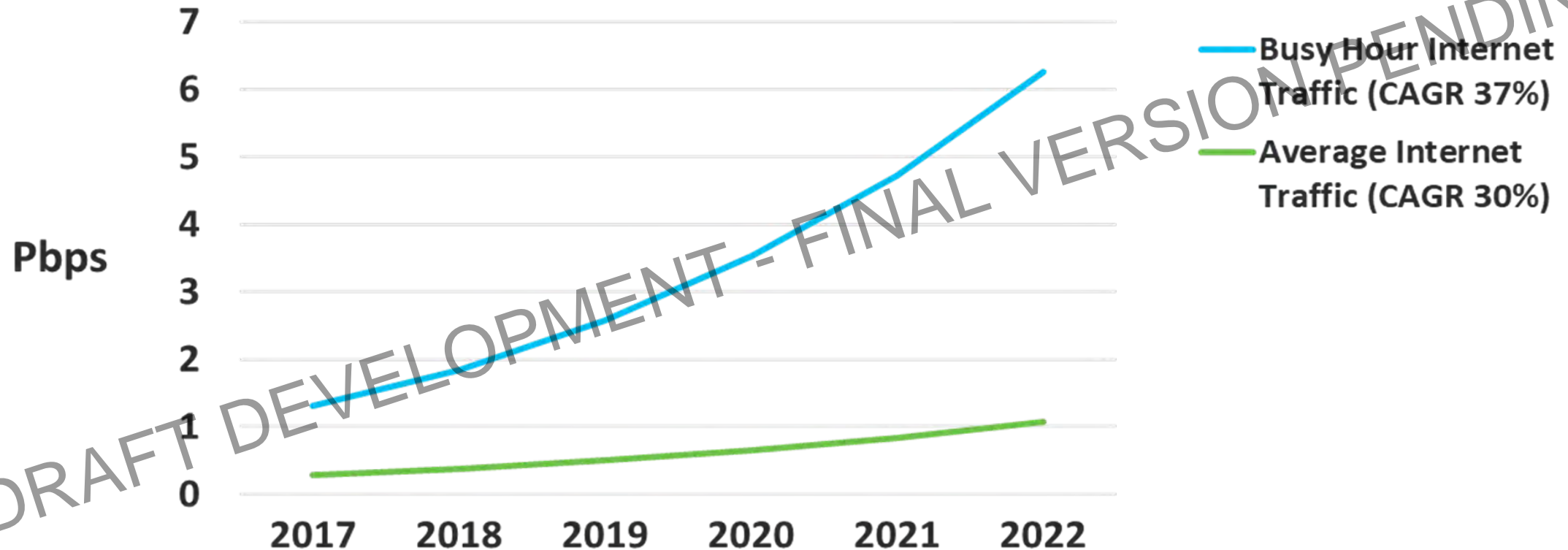
Latin America			
(Mb/s)	2017	2022	CAGR
Fixed Broadband	11.7	28.1	19.2%
Wi-Fi	9.0	16.8	13.3%
Cellular	4.9	17.7	29.3%

Middle East & Africa			
(Mb/s)	2017	2022	CAGR
Fixed Broadband	7.8	20.2	21.0%
Wi-Fi	6.2	11.2	12.6%
Cellular	4.4	15.3	28.3%

Asia Pacific			
(Mb/s)	2017	2022	CAGR
Fixed Broadband	46.2	98.8	16.4%
Wi-Fi	26.7	63.3	18.8%
Cellular	10.6	28.8	22.1%

Source: Cisco VNI Forecast Update, http://www.ieee802.org/3/ad_hoc/bwa2/public/calls/19_0624/nowell_bwa_01_190624.pdf

GLOBAL INTERNET TRAFFIC BUSY-HOUR VS AVERAGE HOUR

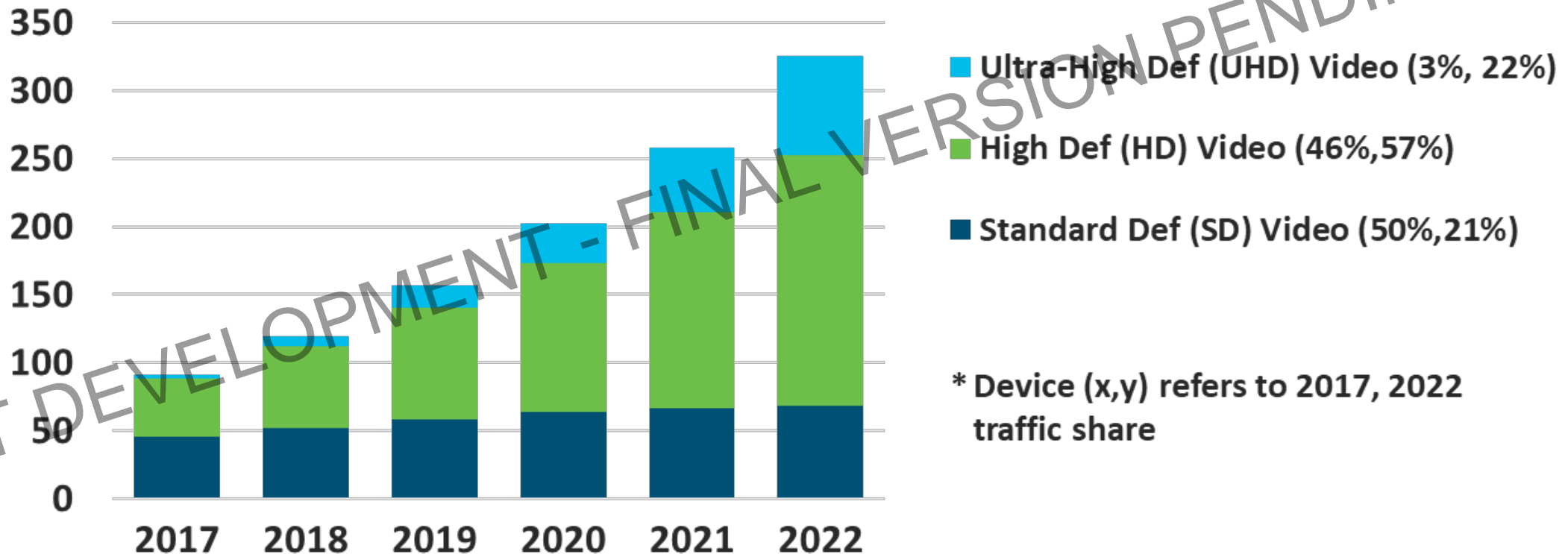


Source: Cisco VNI Forecast Update, http://www.ieee802.org/3/ad_hoc/bwa2/public/calls/19_0624/nowell_bwa_01_190624.pdf

IMPACT OF “DEFINITION” ON IP VIDEO GROWTH

29% CAGR
2017–2022

Exabytes
per
Month

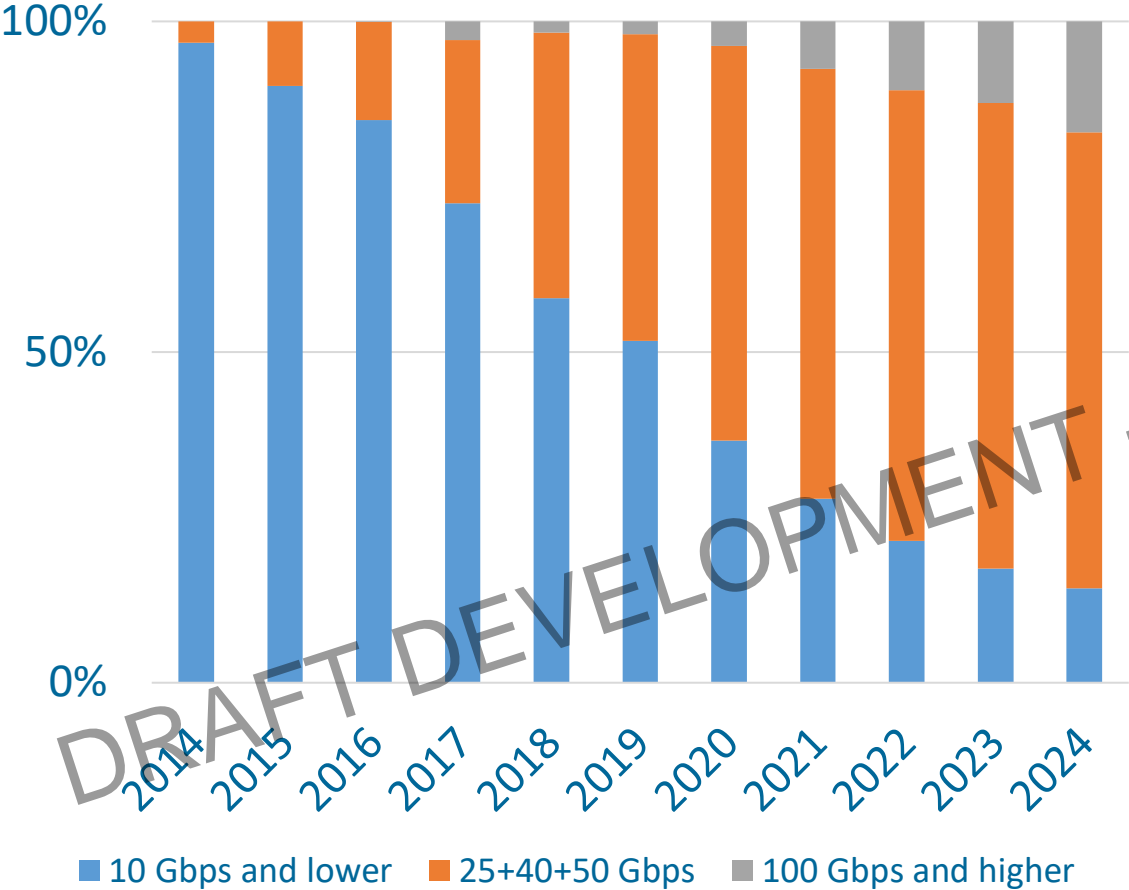


Growth in the adoption of HD and UHD dominate IP video traffic

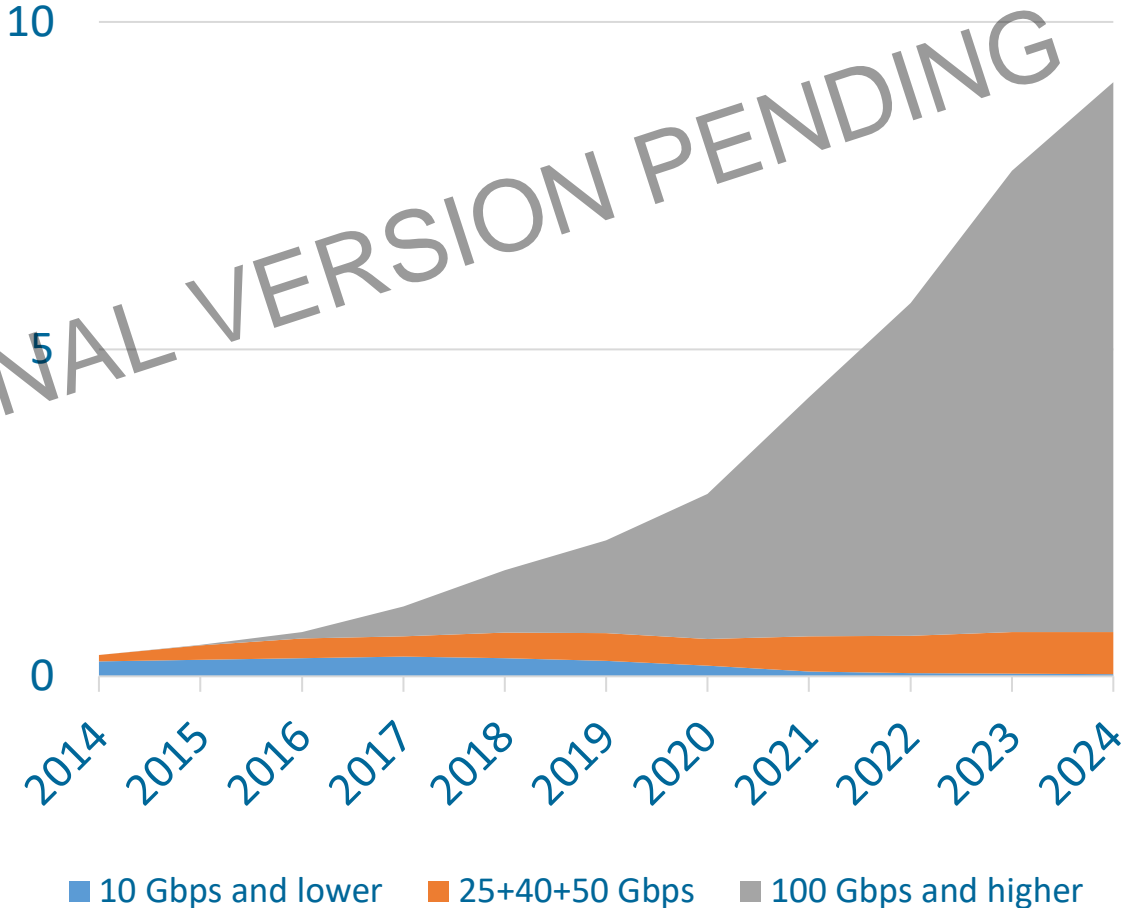
Source: Cisco VNI Forecast Update, http://www.ieee802.org/3/ad_hoc/bwa2/public/calls/19_0624/nowell_bwa_01_190624.pdf

DATA CENTER CAPACITY CONTINUES TO GROW

Enterprise / Cloud Server Unit Shipments*



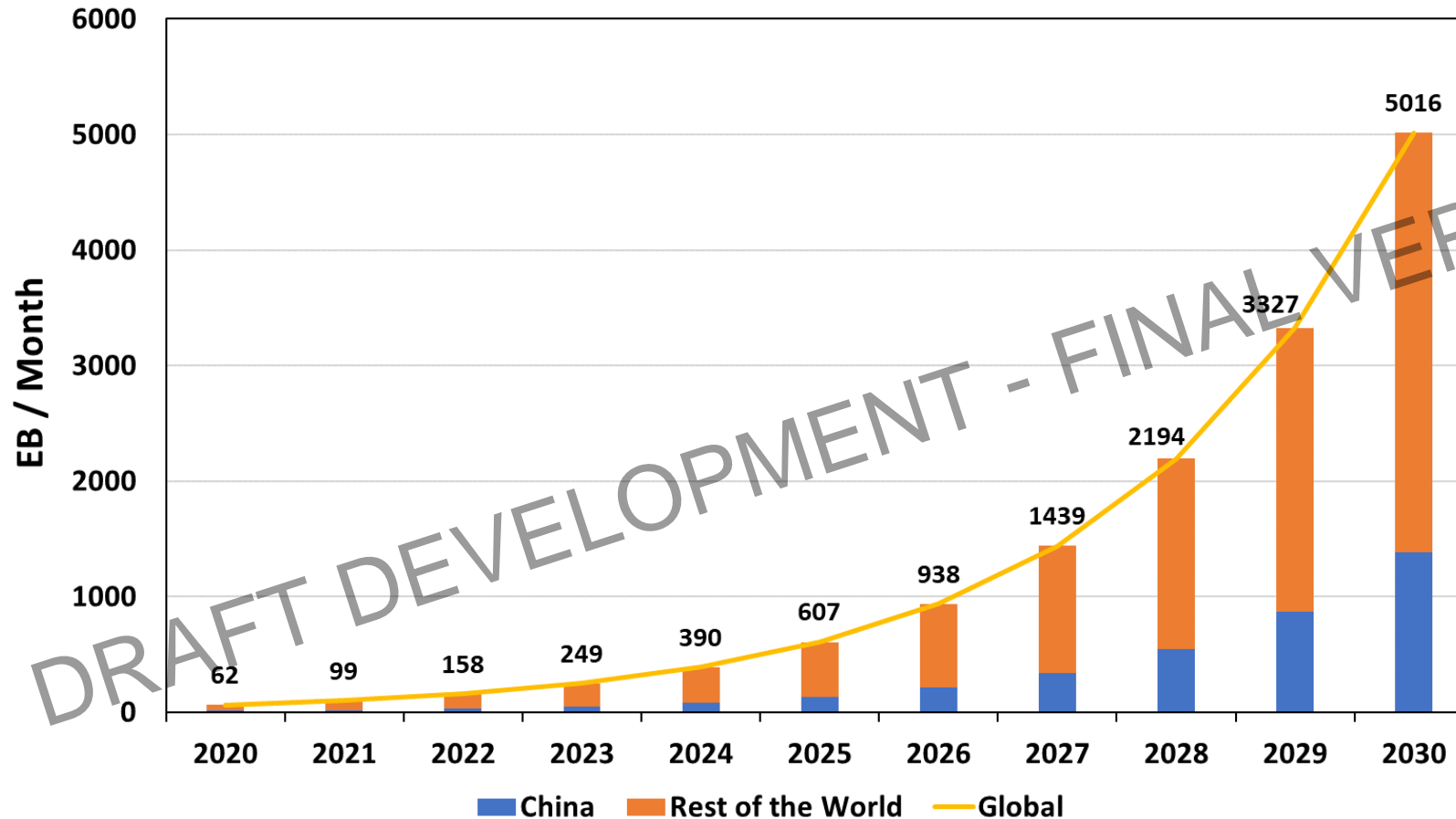
Switch Capacity Shipments in Eb/s**



* Percent of annual server shipments categorized by speed of the attached controllers and adapters
** Annual port capacity shipped on Data Center Ethernet Switches measured in exabits per second



ESTIMATION OF MOBILE TRAFFIC



Global mobile traffic is exponential and may even be underestimated

Source: Report ITU-R M.2370-0: IMT traffic estimates for the years 2020 to 2030, <https://www.itu.int/pub/R-REP-M.2370-2015>