

# Proposed Objectives for Parallel PMDs

Robert Lingle, Jr. (OFS), Mabud Choudhury (OFS)

Beyond 400G Study Group

March 15, 2021

IEEE 802 Plenary Week

# Supporters

- Ramana Murty – Broadcom
- Vipul Bhatt – II-VI
- James Young – CommScope
- Tom Mitcheltree - USConec
- Earl Parsons – CommScope
- David Malicoat – Senko
- Jose Castro – Panduit
- Ali Ghiasi – Ghiasi Quantum
- David Piehler – Dell EMC
- Tom Palkert - Macom
- Kent Lusted – Intel

# Overview

- Traditional benefits of VCSELs+MMF for lower cost & power persist
- MMF more tolerant to dust at connectors
- 100G electrical lanes may persist for a long time
- 100G VCSELs will be high volume, lower-cost parts over the next five years, including with direct drive, finding multiple short-reach applications
- They will be attractive for fiber-to-the-server/GPU with breakout cabling
- 100G VCSELs will be attractive to build 100/200/400/800 GbE parallel links
- Propose to adopt several objectives in B400G Study Group

# Summary of market segments

- Some hyperscale DC users in North America want parallel MMF links with 100G VCSELs over up to 50m lengths for server/GPU/memory attachment
  - With 50 & 100G lanes, server attachment is essentially a breakout application from an octal switch port: 400GBASE-SR8 & 800GBASE-SR8
  - With 50G lanes, also switch links up to 70m using 400GBASE-SR8 as dual-200G
- SR4 links with 4-pair MMF cabling widely used in China Big Cloud DCs
  - 100m reach covers up to 2/3 of switch links, per discussion in 802.3db
  - Current roadmap from 100 to 200 to 400GBASE-SR4 over 4-pair cable
  - Application of 8-pair cabling under study as roadmaps evolve
- The Large Enterprise DC market globally has a large installed base of 4-pair MMF cabling built around traditional 100m SR reach targets
  - 4:1 breakout is useful in these applications
  - Current roadmap from 40 to 100 to 200 to 400GBASE-SR4 & SR4.2
  - 400GBASE-SR4.2 deploying in 2021 in some large enterprises

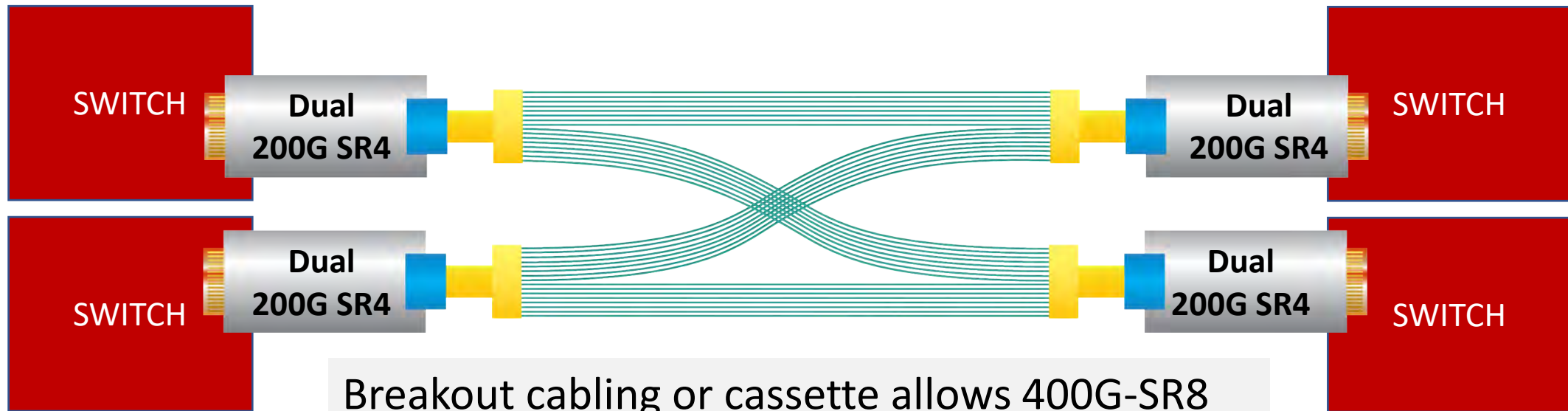
# Applications in the hyperscale datacenter

- 50G lanes over parallel MMF in Clause 138 & 150 (802.3cd & cm complete)
  - 50/100/200/400GBASE-SR1/2/4/8 over 100m
  - Switch links: 400GBASE-SR8 modules used for 2x200G and 1x400G
  - Server attachment: 50G links broken out from 4x50G and 8x50G ports
    - See diagrams in following slides
- 100G lane MMF standards in Clause 167 (802.3db in progress)
  - 100/200/400GBASE-VR1/2/4 over 50m, cost-optimized server attachment in North America hyperscale & short switch links in China cloud
  - 100/200/400GBASE-SR1/2/4 over 100m switch links for large enterprise & China cloud
  - Note: 100G-VR and 100G-SR will interoperate over 50m
- Propose adding 800GBASE-VR8/SR8 in B400G
  - Simple extension of VR/SR PMDs in 802.3db to 800G for low-cost switch links w/ defined MDI
  - VR8 used for 8 x 100G links from octal switch port to servers

# Applications of 400GBASE-SR8 PMD in switch links

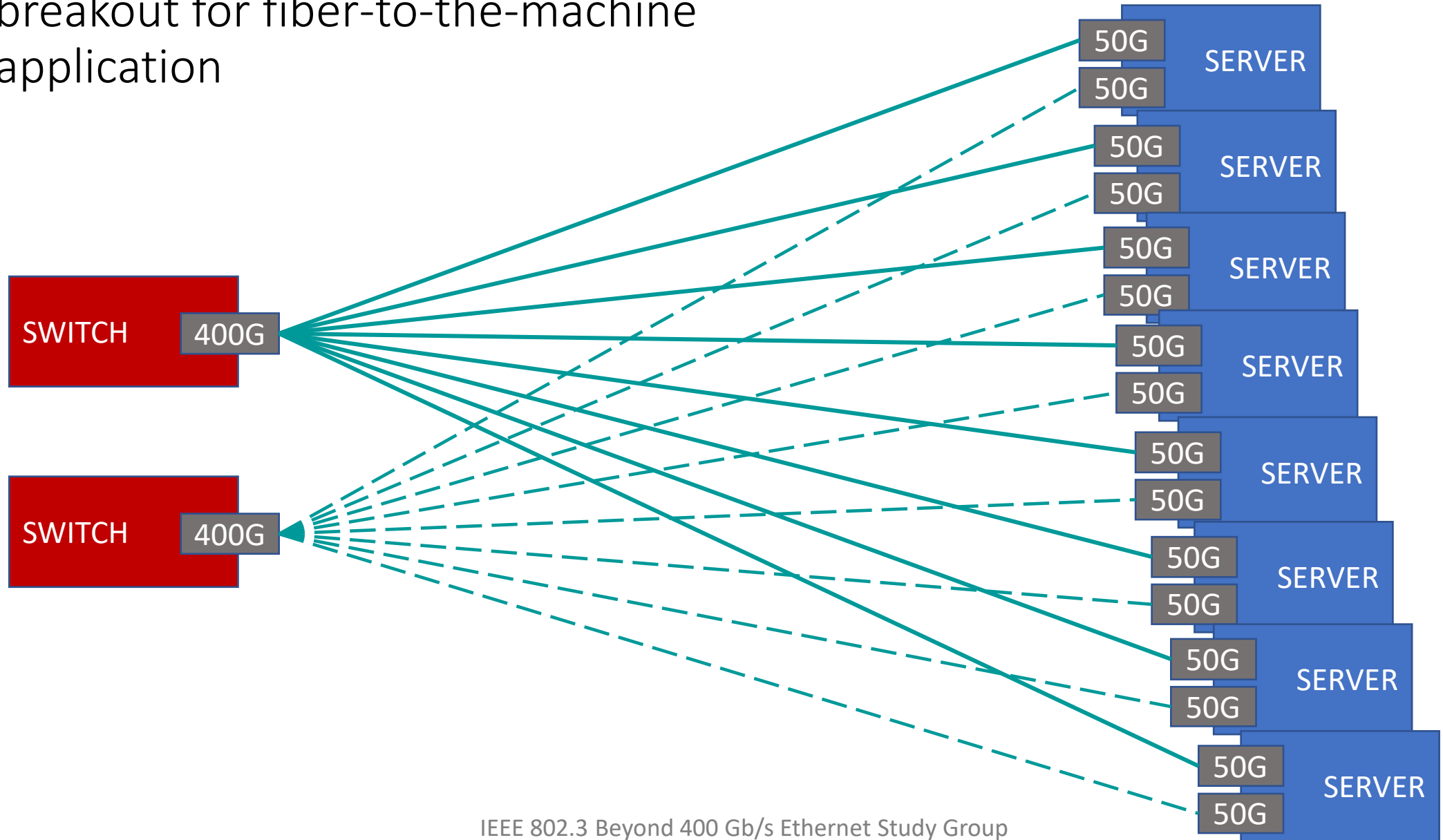


SR8 module shown in 400G point-to-point link



Breakout cabling or cassette allows 400G-SR8 module to function as a dual 200G-SR4 module

# Use of 400GBASE-SR8 modules for 50G breakout for fiber-to-the-machine application



# Key reach scales for MMF links in some China big cloud datacenter architectures

See, for example, the discussion of reach in minutes of November 5, 2020 TF Telephonic Interim

[https://www.ieee802.org/3/db/public/November-05-2020/meeting\\_minutes\\_3db\\_01\\_110520.pdf](https://www.ieee802.org/3/db/public/November-05-2020/meeting_minutes_3db_01_110520.pdf)

**100m reach  
within a  
building**

**50m reach within a floor**



# Applications in the hyperscale datacenter

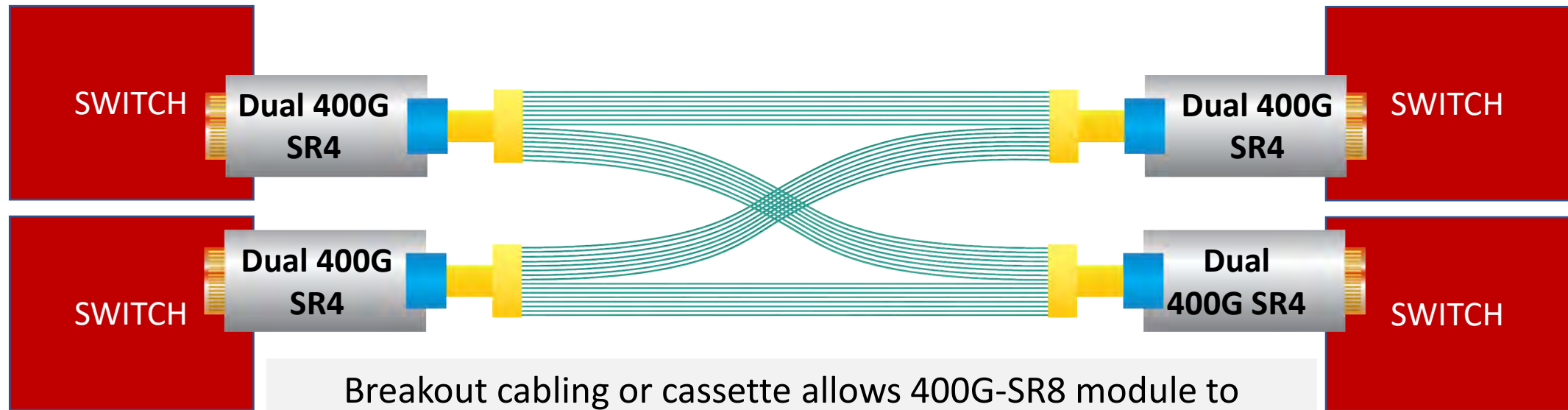
- 50G lanes over parallel MMF in Clause 138 & 150 (802.3cd & cm complete)
  - 50/100/200/400GBASE-SR1/2/4/8 over 100m
  - Switch links: 400GBASE-SR8 modules used for 2x200G and 1x400G
  - Server attachment: 50G links broken out from 4x50G and 8x50G ports
    - See diagrams in following slides

- 100G lane MMF standards in Clause 167 (802.3db in progress)
  - 100/200/400GBASE-VR1/2/4 over 50m, cost-optimized server attachment in North America hyperscale & short switch links in China cloud
  - 100/200/400GBASE-SR1/2/4 over 100m switch links for large enterprise & China cloud
  - Note: 100G-VR and 100G-SR will interoperate over 50m
- Propose adding 800GBASE-VR8/SR8 in B400G
  - Simple extension of VR/SR PMDs in 802.3db to 800G for low-cost switch links w/ defined MDI
  - VR8 used for 8 x 100G links from octal switch port to servers
  - See diagrams in following slides

# Application of proposed 800GBASE-VR8/SR8 PMDs in switch links

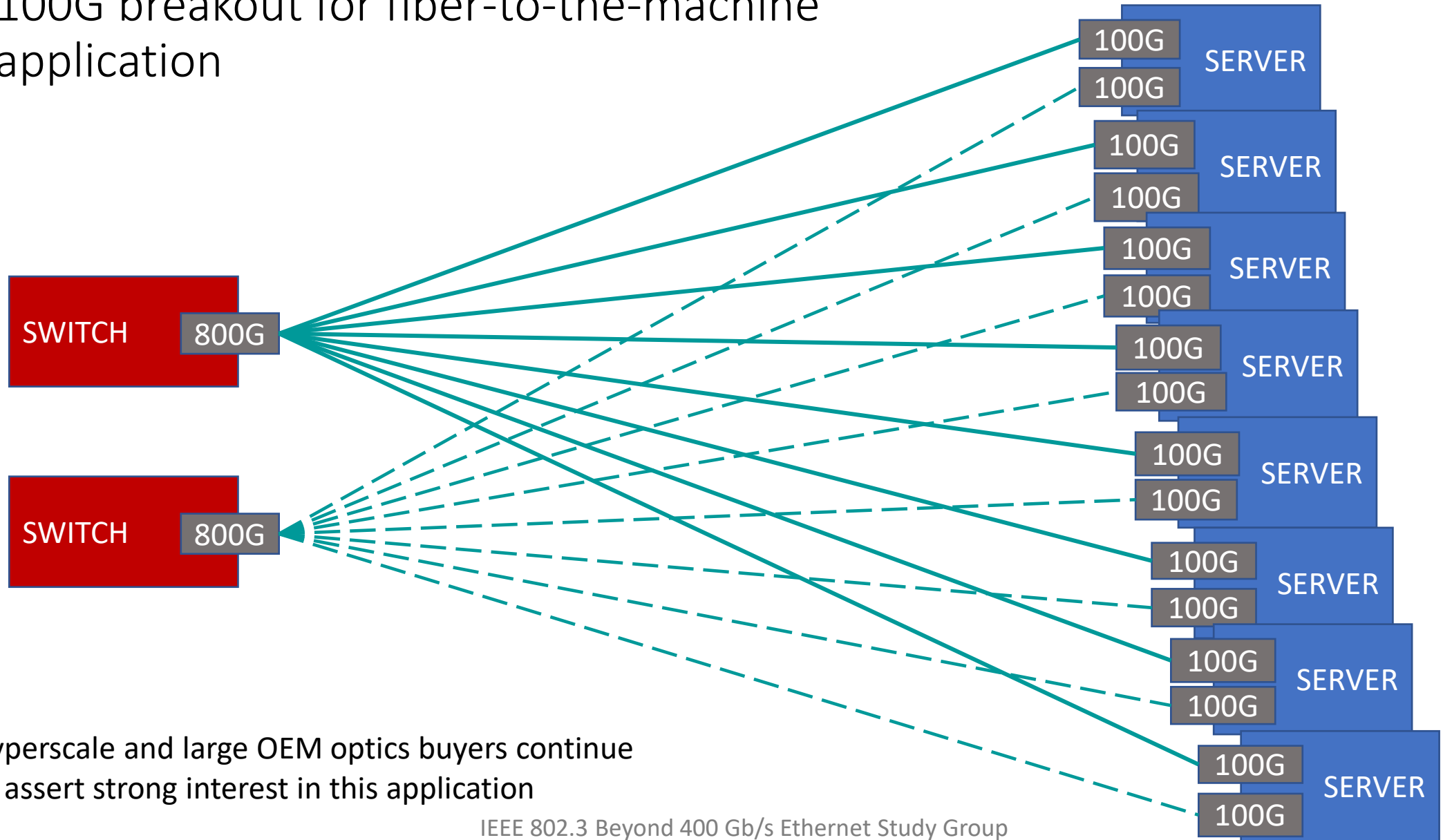


VR8 module shown in 800G point-to-point link over 50m



Breakout cabling or cassette allows 400G-SR8 module to function as a dual 400G-SR4 module up to 100m (potentially applicable to non-hyperscale datacenters)

# Use of 400GBASE-VR8 modules for 100G breakout for fiber-to-the-machine application



Hyperscale and large OEM optics buyers continue to assert strong interest in this application

# In case short-reach 1.6 Tb/s links are needed ...

- It is readily conceivable to build a 1600GBASE-VR8.2 pluggable for 8-pair MMF cabling using the technology developed in IEEE P802.3cm and .3db
  - This would require 16 Tx & 16 Rx, form factor TBD – less desirable
  - Use 8-pair cabling and connectors already in use
  - 850 and 910nm VCSELs
  - By comparison, a 1600GBASE-DR8.2 could use 8 Tx at 200G (optical)
- 1600GBASE-VR16 could be easily adapted for on-board or co-packaged implementations if applications arose
  - 850nm VCSELs only
  - Use of thin or multi-core fibers to save space

## SR4 options for installed based of 4-pair MMF cabling

Lane Rate	PMD for 4p cabling	MAC data rate
10G	40GBASE-SR4	40G
25G	100GBASE-SR4	100G
50G	200GBASE-SR4	200G
50G	400GBASE-SR4.2	400G
100G	400GBASE-SR4	400G
100G	?800GBASE-SR4.2?	800G

- For 50G lanes, 100m over installed base 4-pair cable by 400GBASE-SR4.2
  - SR4.2 refers to use of 850 & 910nm wavelengths
  - Technology advance in many large enterprises stalled by COVID in 2020
  - Industry M&A, reorgs slowed progress; several large OEMs plan to launch 2021
- For 100G lanes, 400GBASE-SR4 under development in 802.3db
- For 100G lanes, 800GBASE-SR4.2 possible for the 4-pair MMF cable base
  - 200G electrical may require CPO at the switch, likely leading to long lifetime of 100G SerDes ASICs in large enterprise DCs. 800 GbE likely to be deployed on this platform
  - Technical feasibility can be demonstrated today for 70/100m over OM4/5
  - Further work needed to study technical feasibility for 100m over OM4
  - 200G SR1.2 may be needed as a breakout pair

# Some connectors available for parallel PMDs

Fiber Type	# fiber pairs	Connector	IEC status – mechanical interface standards	PMD Example
MMF	4	MPO-12	Existing standard IEC 61754-7-1	SR4 / SR4.2
MMF	4	Breakout types <sup>1</sup>	SN/SAC: IEC 61754-36, forecast publication date: 2022-12 <sup>2</sup> MDC: IEC 61754-37, forecast publication date: 2022-12 <sup>2</sup>	SR4 / SR4.2
MMF	8	MPO-16	IEC 61754-7-4. Forecast publication date: 2022-02 <sup>2</sup>	SR8 / SR4 <sup>3</sup>
SMF	4	MPO-12	Existing standard IEC 61754-7-1	DR4
SMF	4	Breakout types <sup>1</sup>	SN/SAC: IEC 61754-36, forecast publication date: 2022-12 <sup>2</sup> MDC: IEC 61754-37, forecast publication date: 2022-12 <sup>2</sup>	DR4
SMF	8	MPO-16	IEC 61754-7-4. Forecast publication date: 2022-02 <sup>2</sup>	Did not notice any proposals

Notes: 1) Breakout types include the “SN/SAC” and “MDC”

2) Best understanding is that actual publication date is likely to slip beyond current forecast

3) Use of MPO-16 for 4-pair MMF PMDs will be proposed in 802.3db for the MDI

# 8-pair MMF cable considerations

- The installed base for switch links in hyperscale & large enterprise DCs comprised 4-pair cable with MPO-12 termination
- 400GBASE-SR8 PMD type for use with 8-pair cabling was recommended in IEEE P802.3cm TF by a participant affiliated with a hyperscale DC operator
  - 8-pair fiber cords are readily available
  - MPO-16 APC connectors are shipping & in standardization in IEC, scaling up
- Believe that many 400GBASE-SR8 modules shipped to date have been deployed as dual-200G modules
  - With appropriate connectivity, can preserve trunk cables deployed for SR4
- Low impediments to deploy SR8 for fiber breakout to servers/GPUs

# 8-pair SMF cabling considerations

- Not aware that anyone is proposing this in B400G SG
- Large volumes of G.652.D fiber, often also meeting G.657 for bend insensitivity, have been sold into hyperscale DCs in past 5+ years
- 144f, 288f, and similar, trunk cables are popular in switch fabrics inside one building. Both MPO-12 & MPO-24 trunk terminations utilized.
- Installations deploying PSM4 at 100G up to 500m with 4 pairs per link differ greatly from those focusing on CWDM4 at 100G with 1 pair per link.
- With appropriate connectivity, use of DR8 as dual-400GbE could be mapped to trunk cables deployed for 100GbE PSM4 or 400GbE DR4
- Adapting PSM4 or CWDM4 deployments for 800 GbE with DR8 may add considerable complexity, depending on:
  - greenfield vs. brownfield
  - Link counts → needed fiber counts
  - Appropriate connectivity to map MPO-16 at TRx to MPO-12 & 24 at trunk cables



# Proposed objectives for B400G Study Group

Propose B400G SG should adopt:

- 800 Gb/s operation over 8 pairs of MMF with lengths up to at least 50m (VR8)
- 800 Gb/s operation over 8 pairs of MMF with lengths up to at least 100m (SR8)

Propose B400G SG should further study:

- 200 Gb/s & 800 Gb/s over up to at least 100m over 1 & 4 pairs, respectively, of MMF (SR1.2 & SR4.2)