50G, 100G & 200G SERVER CONNECTIVITY

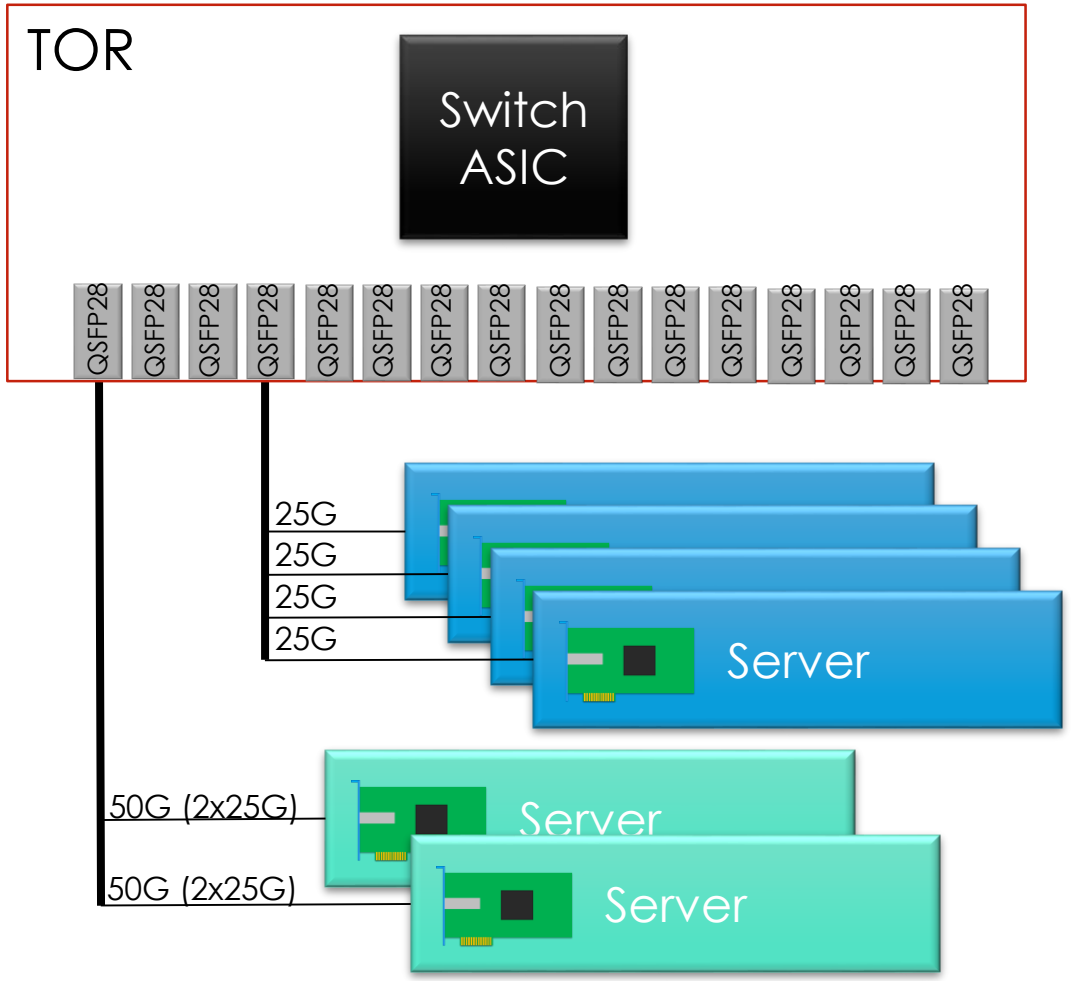
Brad Booth, Microsoft

SUPPORTERS

- Mike Andrewartha, Microsoft
- Phil Sun, Credo
- Silas Li, Precise-ITC
- Scott Sommers, Molex
- Tom Issenhuth, Microsoft
- Rob Stone, Broadcom
- Jeff Twombly, Credo
- Eric Baden, Broadcom
- Bharat Tailor, Semtech

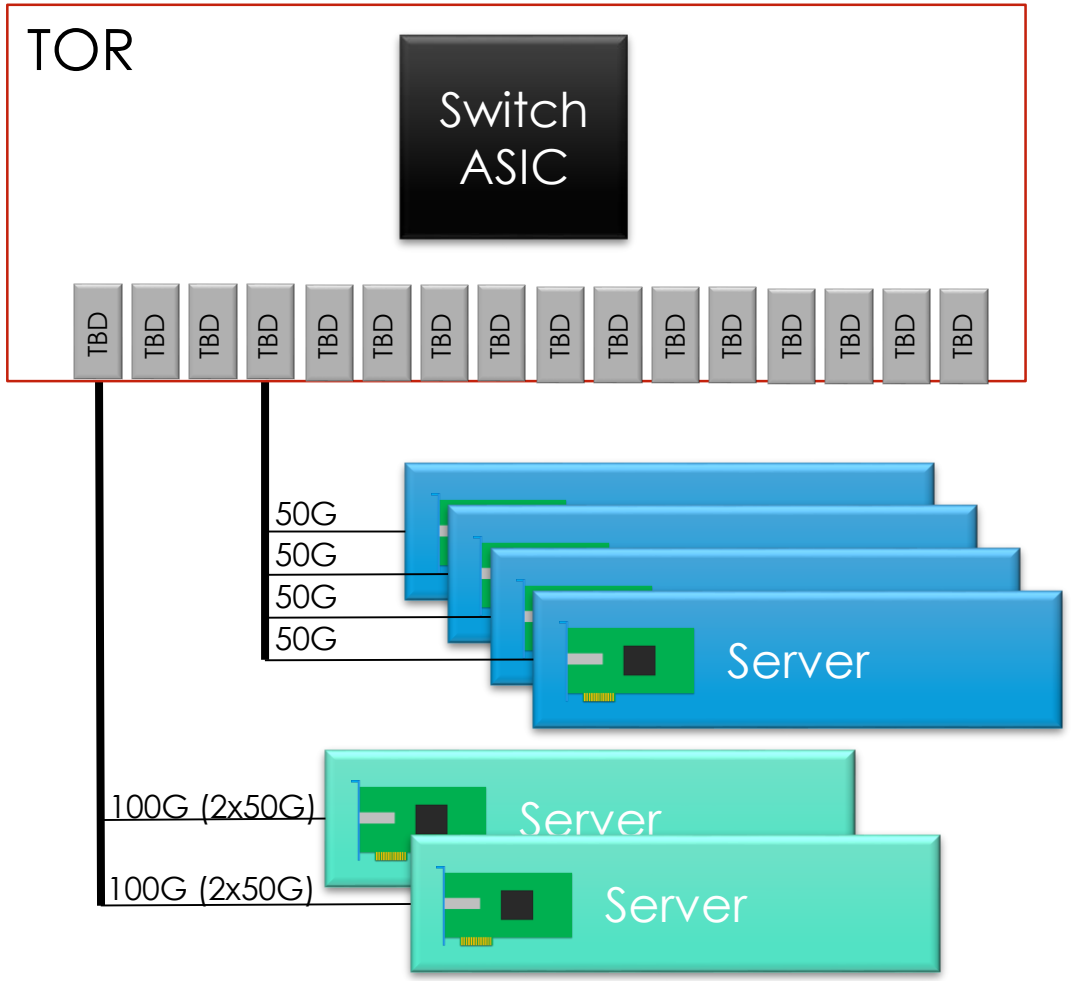
CURRENT TOPOLOGIES

- TOR uplinks are 100G
- 25G server links
  - Single with SFP28, or
  - 4:1 break-out with QSFP28/µQSFP
- 50G server links
  - 25G/50G Ethernet Consortium spec
  - 4:2 break-out with QSFP28/µQSFP
- P802.3by supports:
  - Reach up to 5 meters
  - No FEC, KR-FEC, RS-FEC



POTENTIAL FUTURE TOPOLOGIES

- Uplinks from the TOR are 400G
  - As per booth_400_01a_1113.pdf
- 50G server links
  - Same scenario as with 25G
- 100G server links
  - Very likely given current gen 50G
- 200G server links
  - Interesting with 100G serial links
- TBD module
  - Same module as previous gen?
  - Power requirements DAC vs. optics



THOUGHTS ON 50G, 100G & 200G

- Discussions occurred with datacenter team members on requirements
- Due to timeline, the 100G (2x50G) copper cabling effort is of interest
  - Potential for optics to the server… can there be cost parity?
- Still have a strong requirement for no or very low latency FEC
  - Support bump-in-the-wire technologies
  - Less than 100 ns
- Willing to tolerate a shorter reach to achieve lower latency
  - Up to 2 m on twinax
- 200G MAC-to-MAC above the server is not interesting
  - 200G break-out module is likely to be used

BREAK-OUT

- While typically an implementation issue, do we have the margin at 50G with PAM4?
- 50G PAM4 used in 400G is designed for long reaches
- What happens if we don't use KP4 FEC in short reaches?
- Should the task force consider the implementation concerns with break-out
  - Crosstalk, etc.
  - Do we need an objective?

THANK YOU!