In-row server applications

James Young (CommScope)

April 23, 2020 Ad Hoc meeting

IEEE 802.3 100 Gb/s Wavelength Short Reach PHYs Study Group

Summary

- End User distance requirements for in-row MOR (T1) to servers
  - 30m minimum distance requested (no end user feedback for <30m)
  - Longer distances would support additional applications and increase market appeal (TOR – T1 sw to sw links, Multi-row PODs)
- Several trends combine to favor longer reach for server-attachment, perhaps 50m, longer if possible
  - System level value of High radix switching replacing TOR positioned switches
    - Lower capex, opex and latency
  - Servers including accelerators for AI and ML drive I/O speeds to 50 and 100G
    - http://www.ieee802.org/3/NGMMF/public/Jan18/shen_NGMMF_01_jan18.pdf
- Technically feasible objectives for 100GbE at 70m simulated
  - https://www.osapublishing.org/abstract.cfm?uri=OFC-2020-M3D.5

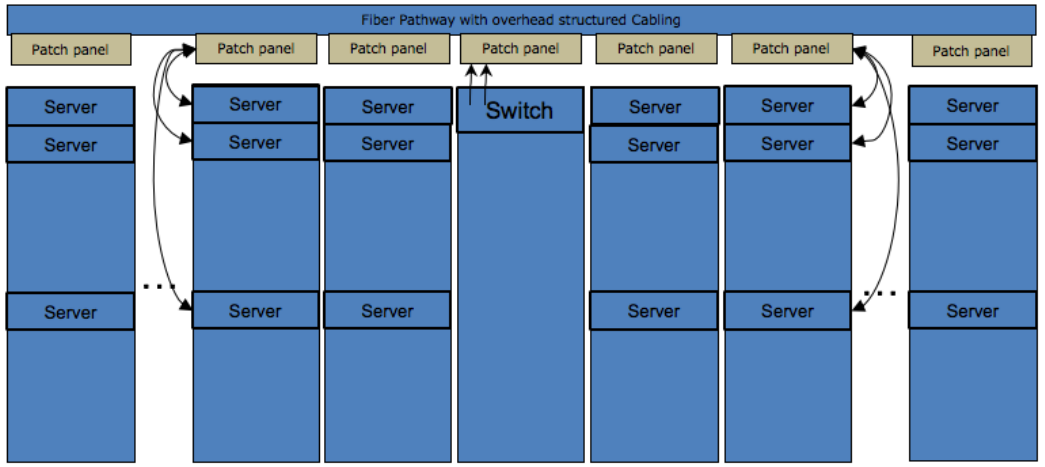
100G 30m to facilitate MOR T1 switch placement

Supports server-row cabling objectives

- Enable pre-installed overhead cabling that supports multiple line rate generations (50/100G) @ 30m
  - Attach to overhead cabling with short cords
  - Repeat installation pattern for all server racks for installation efficiency of ≤ 5 hours for a server row - *Rich Baca (Microsoft)*
- *Chinese ODCC "30m is a reasonable first goal for MMF or AOC transceivers in server interconnects"*

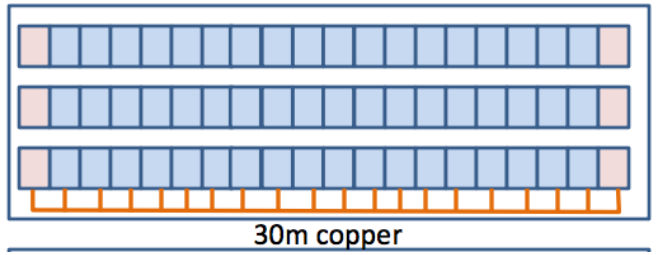
  http://www.ieee802.org/3/cfi/1119_1/CFI_01_1119.pdf
- *30m cable length meets the market's need*
  - *Basic DC distances haven't changed*
  - *Server Row applications still 30m*

  http://www.ieee802.org/3/NGBASET/public/nov12/wagner_01_1112_ngbt.pdf

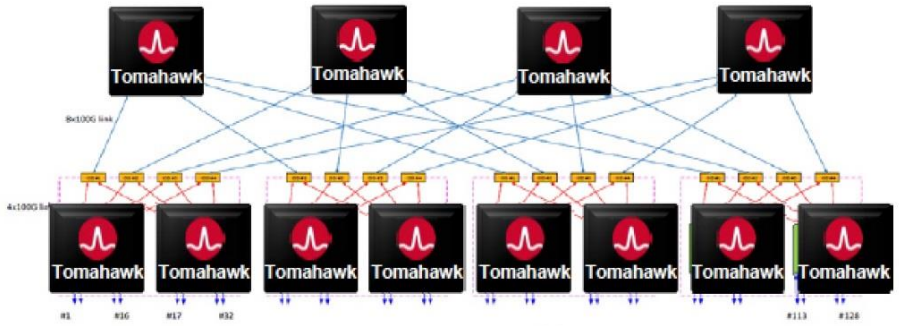


- Typical server row 16 – 20 cabinets
- Cabinets arrive on site with servers installed
- Overhead cable is pre-installed with pathway
- Simple patching from server to overhead patch panel

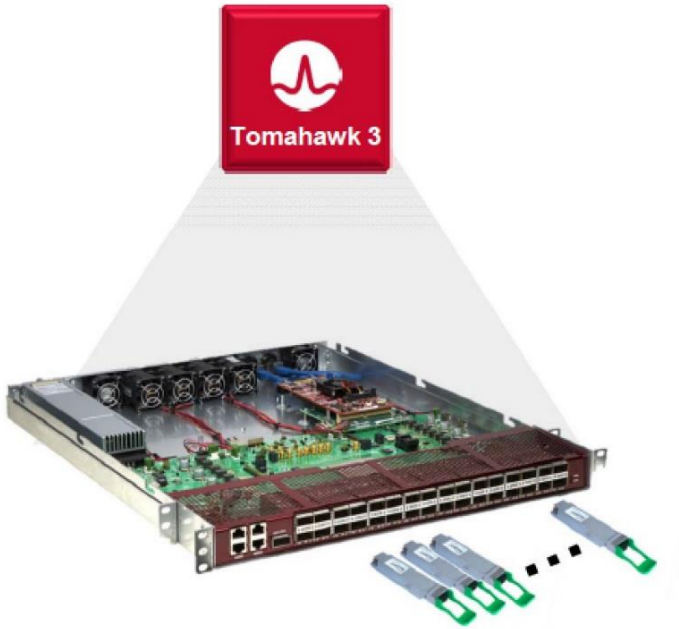
Deployment Topologies



End of Row (or Middle of Row)
- Switch at end of row
- Structured cabling along row, 30m max. channel
- Some use cross-connect on switch end
- Switching at each end for redundancy
- Fiber typically for inter-switch uplink, although copper can be used



: Facebook, OCP

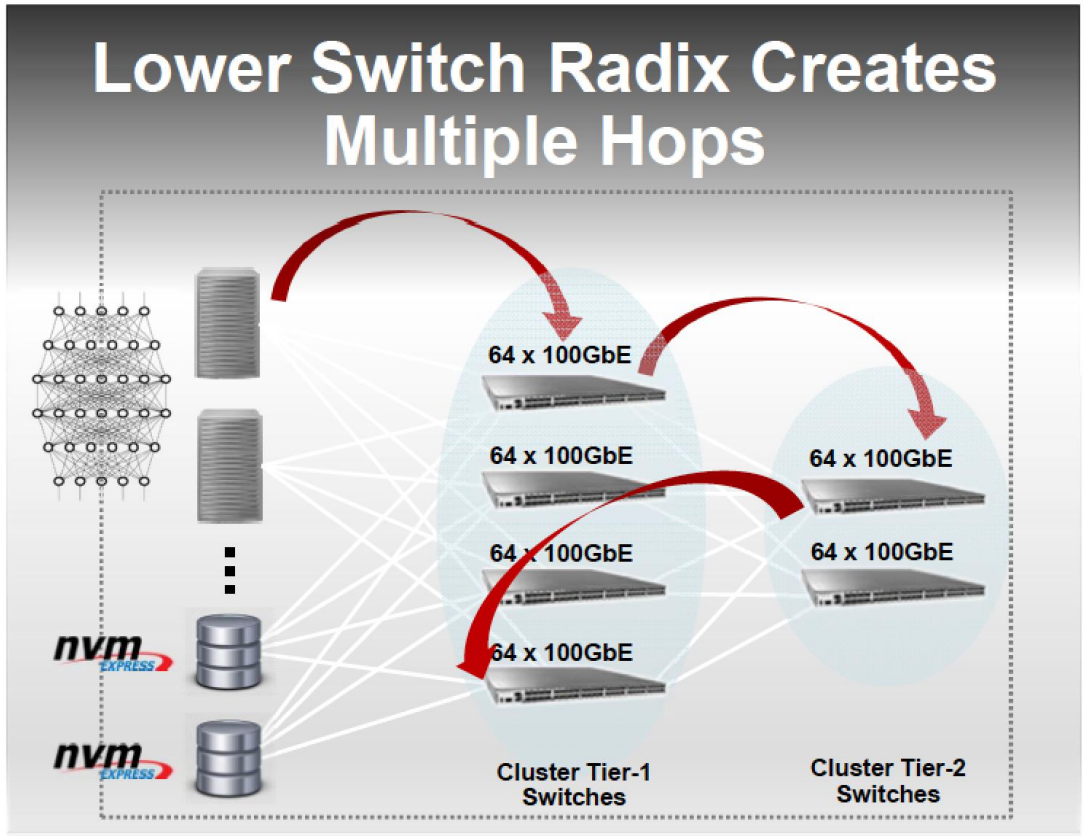


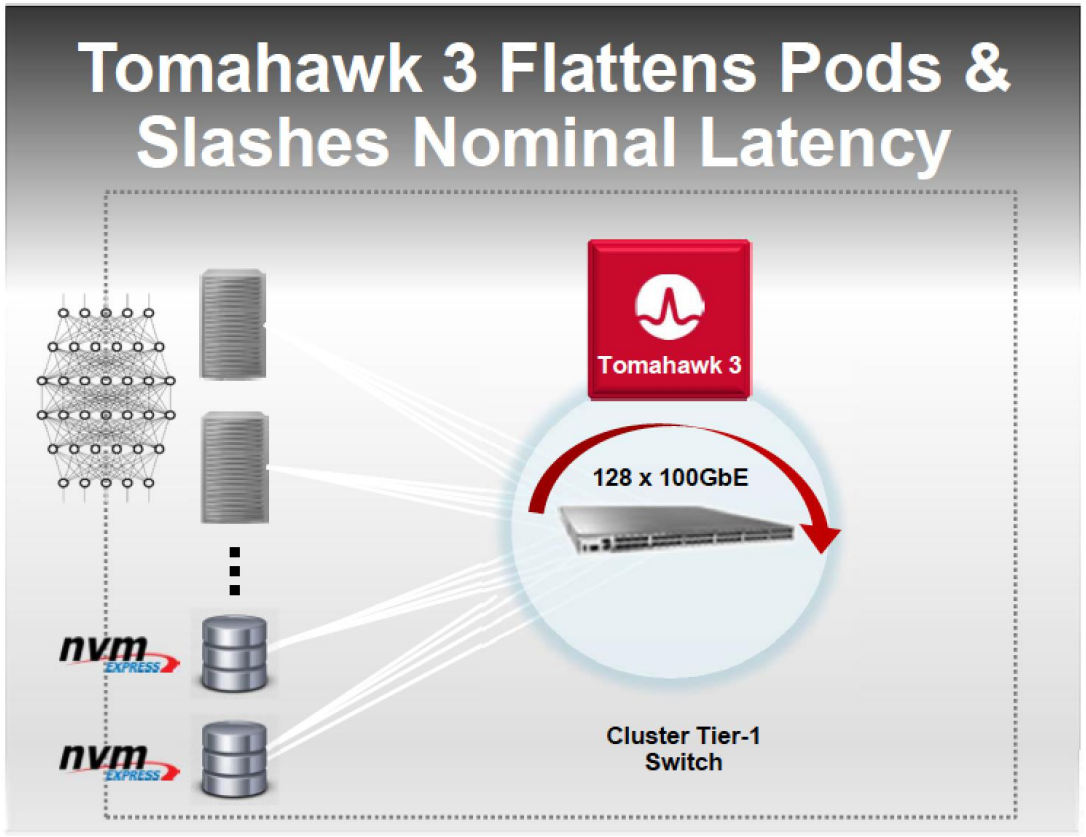
	FB Backpack	Next Gen
Capacity	128x100GbE	32x400GbE / 128x100GbE
Front panel	128x QSFP28	32x QSFP-DD or OSFP
Size	8U Chassis	1U Fixed
# Switch Chips	12 x Tomahawk	1 x Tomahawk 3

75% reduction in power and 85% reduction in system cost

Increasing Radix decreases $$ and lowers latency

Source Broadcom





Source Broadcom