25GBASE-T Call For Interest Consensus

IEEE 802 November 2014 Plenary, San Antonio, TX

Contributors and Supporters

- David Chalupsky, Intel
- George Zimmerman
  - CME, Aquantia, Commmscope
- Kamal Dalmia, Aquantia
- Jonathan King, Finisar
- Tom Souvignier, Broadcom
- Kent Lusted, Intel
- John D'Ambrosia, Dell
- Theo Brillhart, Fluke Networks
- Derek Cassidy, BT
- Bob Wagner, Panduit
- Chris DiMinico,
  - MC Communications, Panduit
- Valerie Maguire, Siemon
- Vineet Salunke, Cisco
- Rich Mellitz, Intel
- Ron Nordin, Panduit
- Peter Jones, Cisco
- Thuyen Dinh, Pulse
- Steve Carlson,
  - High Speed Design
- Yong Kim, Broadcom
- Martin Rossbach, Nexans
- Alan Flatman,
  - LAN Technologies
- Paul Vanderlaan, Nexans
- Andrew Jimenez, Anixter
- Yakov Belopolsky
  - Bel Stewart Connector
- Chris Cole, Finisar
- Matt Brown, APM
- Victor Renteria
  - Bel Stewart Connector
- Richard Mei, Commscope
- Wayne Larsen, Commscope
- German Feyh, Broadcom
- Rich Hernandez, Dell
- Shadi AbuGhazaleh, Hubbell
- Brett McClellan, Marvell
- Jeff Poulsen, Leviton
- Paul Langner, Aquantia
- William Lo, Marvell
- Anna An, FIT
- David Ofelt, Juniper
- Jerry Chiang, FIT
- Peter Wu, Marvell
- Scott Kipp, Brocade
- Amrik Bains, Cisco
- Mark Nowell, Cisco
- Ghani Abbas, Ericsson
- Jon Lewis, Dell
- Pete Cibula, Intel
- Allen SJ Wang, Lenovo
- Brad Booth, Microsoft

Objectives

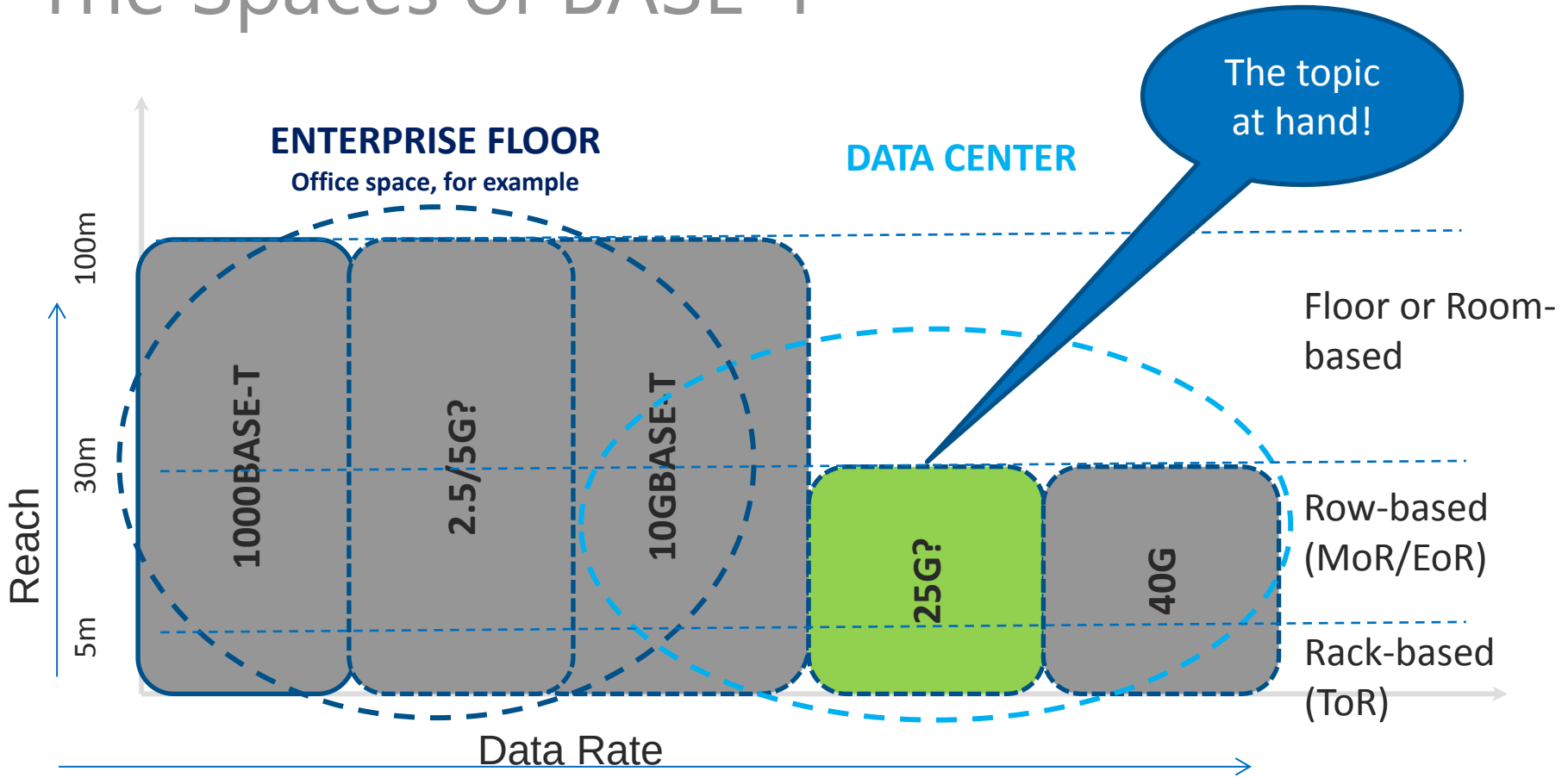
- To gauge the interest in starting a study group to investigate a 25GBASE-T project

- We do not need to:
  - Fully explore the problem
  - Debate strengths and weaknesses of solutions
  - Choose a solution
  - Create a PAR or 5 Criteria
  - Create a standard
- Anyone in the room may vote or speak

Agenda

- 25GBASE-T Market Drivers
- 25GBASE-T Technical Feasibility
- 25GBASE-T Why Now?
- Straw Polls

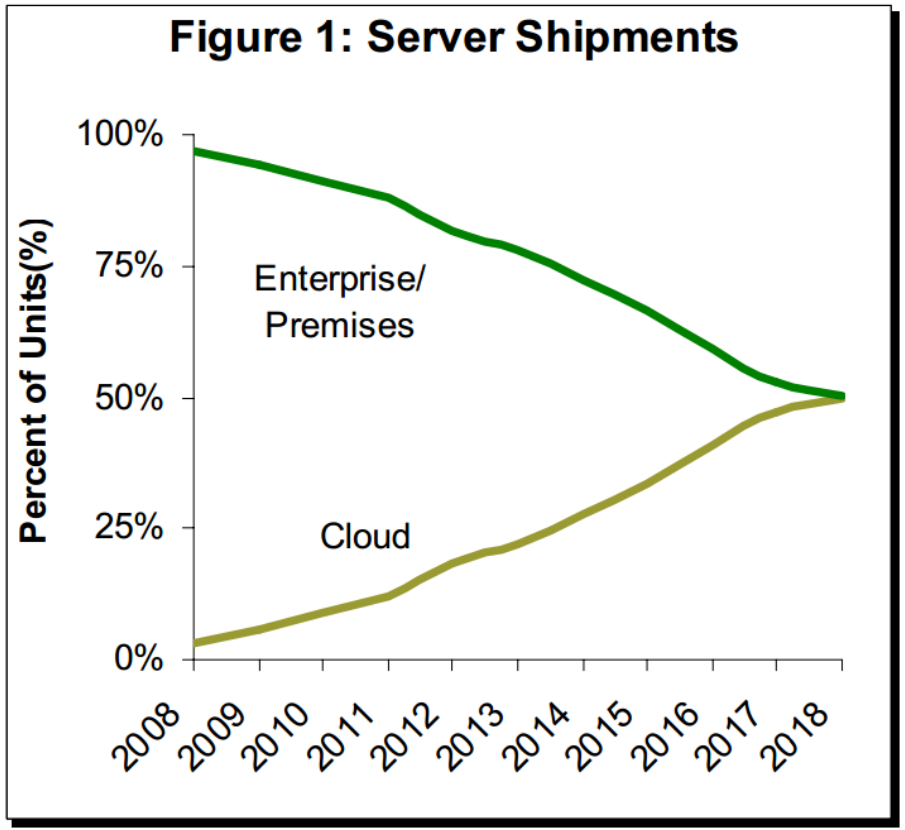
The Spaces of BASE-T



25GBASE-T Market Drivers

- Cloud vs. Enterprise Servers
- Update on server port speed forecast
- Media mix
- Topology
- Cost Optimization

Server Deployment in Cloud vs. Enterprise



Source: Dell'Oro Controller and Adapter Report, July 2014

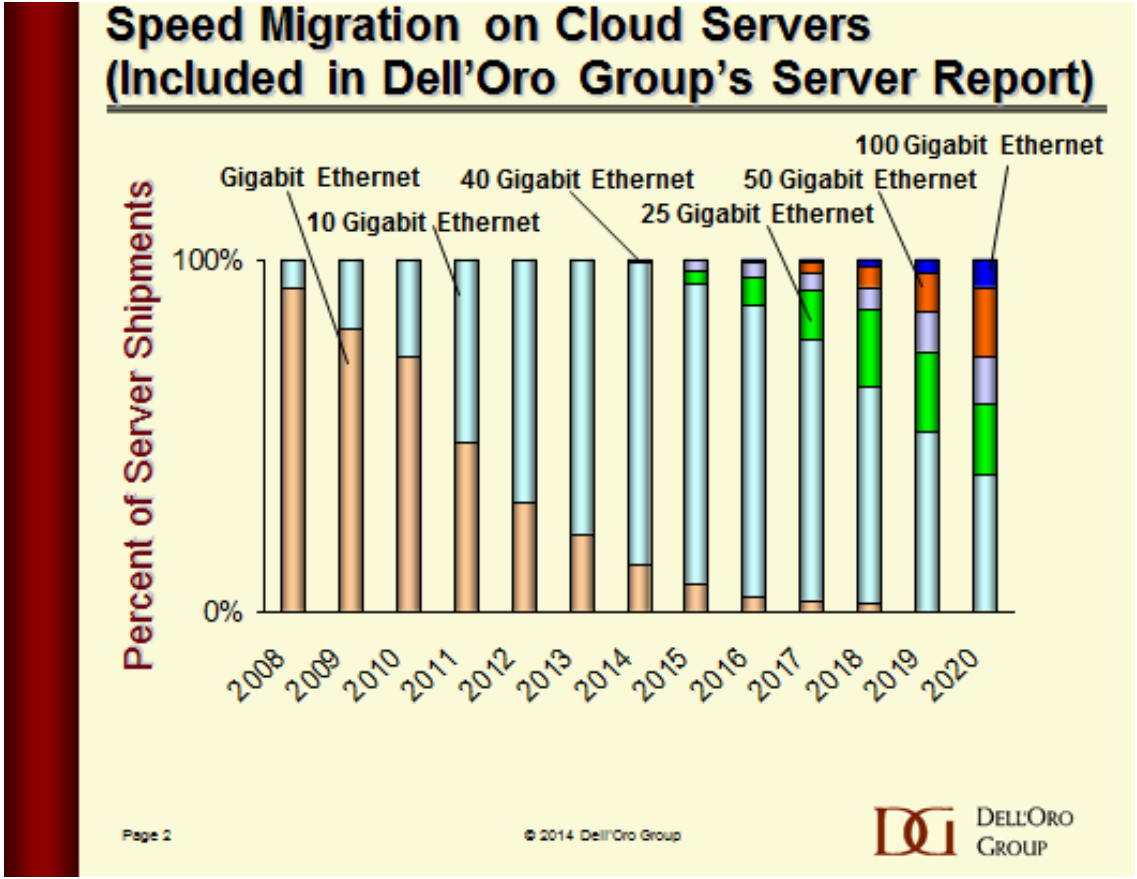
Cloud: Computing *is the business*

- ~1/4 of Server units in 2014
- Growing fast, will stabilize
- A small number of entities comprise the majority of this market

Enterprise IT: Computing *supports the business*

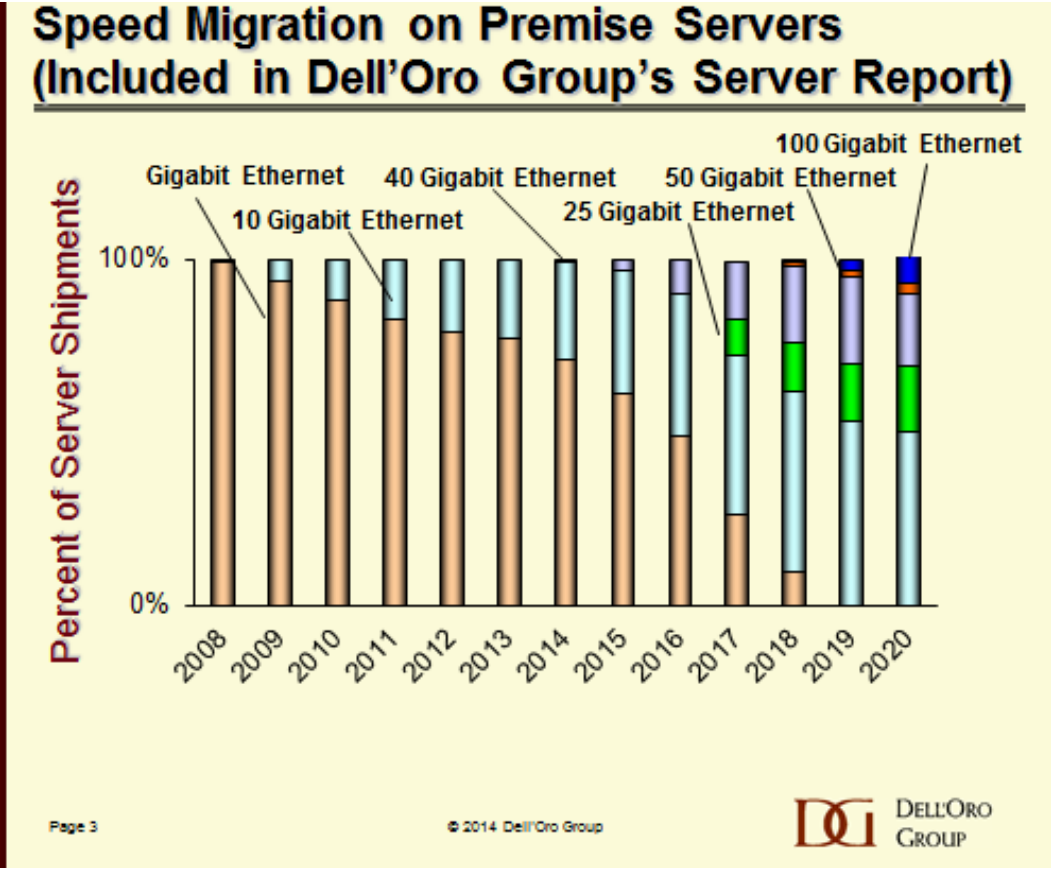
- ~39,000 US companies with >100 employees, typically 100-1000 servers
- Largely follow TIA-942 Telecommunications Infrastructure Standard for Data Centers and deploy a variety of access switch placements including MoR and EoR.

Speed Migration on ***Cloud Servers***



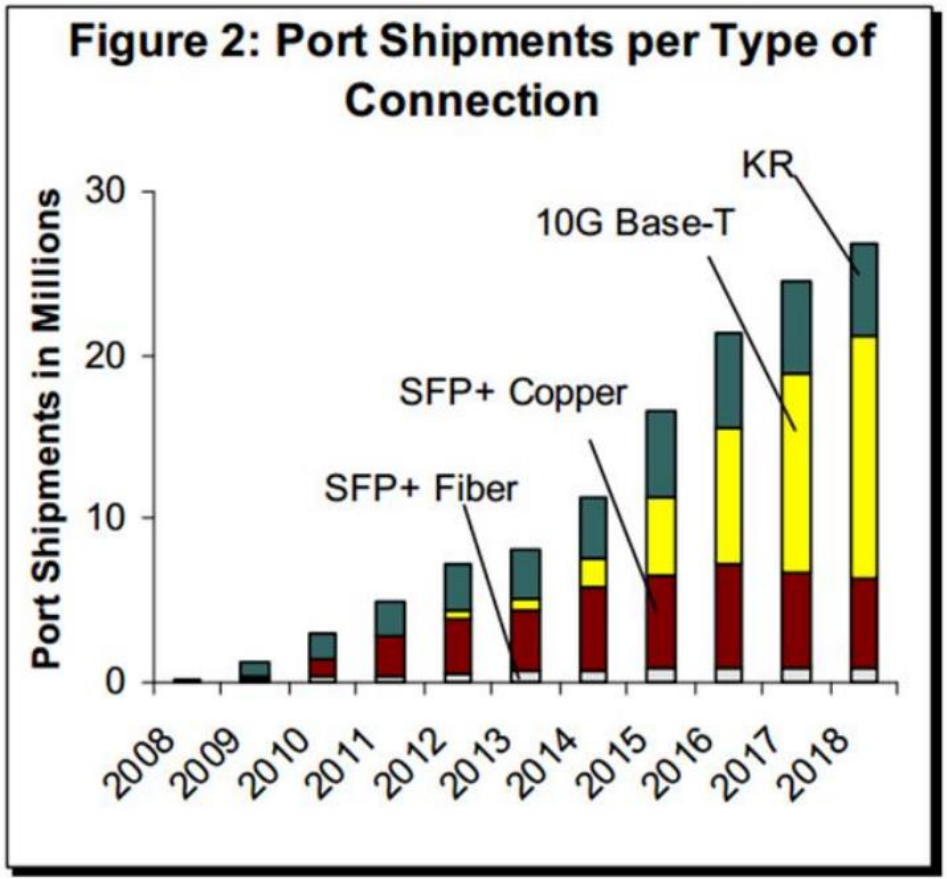
- Transition to 10G nearly complete
- Early adoption of 40G, but 25G/50G expected to overtake
- Standardization not imperative
  - Engineered systems

Speed Migration on ***Premise Servers*** *i.e. Enterprise*



- 1000BASE-T still the majority of ports
- 10G growing in enterprise server & data center
- 40G starts deployment now; significant future forecast
- Investment depends upon the economy
- Standardization important
  - 25G considered for deployment once standardized, and multi-vendor, off the shelf hardware available

10G Server Port Type Mix



- 10G Early Adopters on Fiber, Backplane, and Twinax
- 10GBASE-T forecast grows as 1G→ 10G transition progresses in Enterprise data center

Source: Dell'Oro Controller and Adapter Forecast Summary – January 2014

Diversity of DC architectures

from king_25GE_02a_0914.pdf

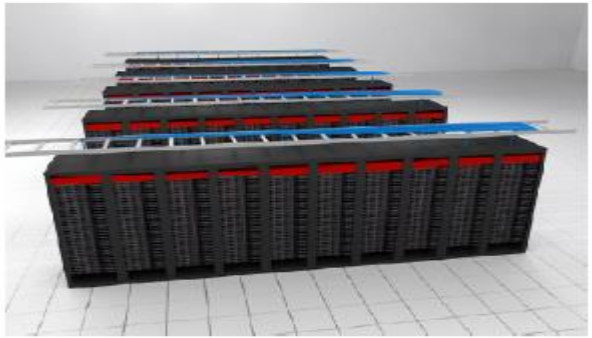
"There is no single end-all cabling configuration for every data center"[a]

- ToR is not sufficient for all applications
  - A single 1U switch can connect to multiple racks of servers[c]
  - Modular switches can connect to 1000's of servers over many tens of racks[c]
    - A 3m to 5 m reach PHY addresses just intra-rack cabling
- A longer reach PHY enables cabinet to cabinet, Middle of Row (MoR) and End of Row (EoR) data center architectures
- Each topology (ToR, MoR, EoR) has pros & cons, and exists in the market today for very sound reasons:
  - Cost, switch port use efficiency, thermal management, maintenance, scalability, support of mixed applications, reconfigurability, etc.
  - Abundant supporting material (appendix) ! [a,b,c,d,e,f]
- A 25GBASE-T PHY will have the reach to support a much broader range of DC architectures.

ToR, MoR, EoR Interconnects

Switches
Servers
Interconnects

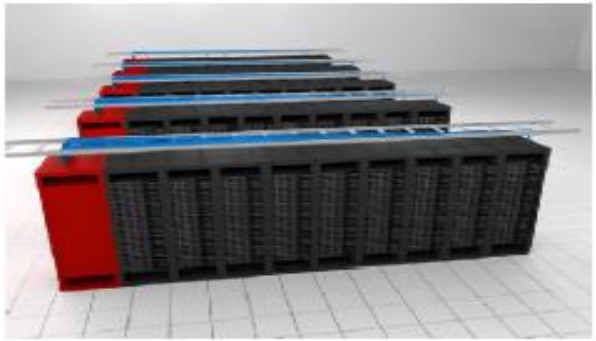
ToR



Intra-rack can be addressed by 25Gb/s twinax copper direct attach

MoR

EoR

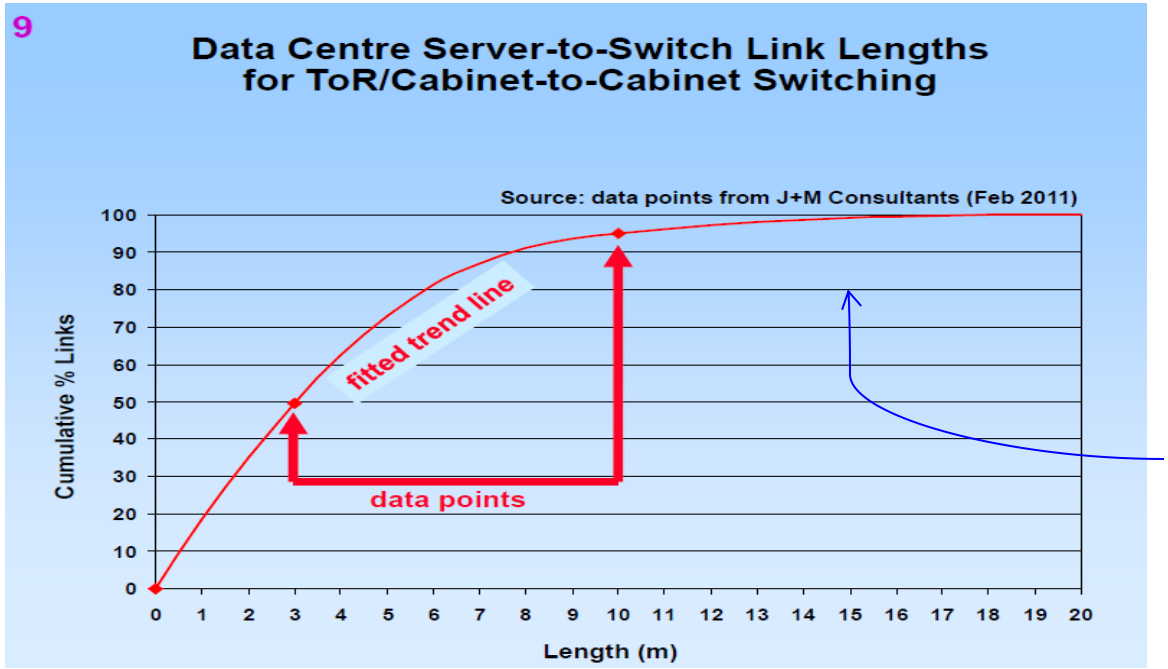


Not addressed by 25Gb/s copper direct attach

Pictures from jimenez_3bq_01_0711.pdf, 802.3bq

ToR link distributions

- Includes cabinet-to-cabinet links
  - Note: slide 8 of *CFI_01_0714*, "The term "TOR" has become synonymous with server access switch, even if it is not located "top of rack", acknowledging that a 3 m reach may not be sufficient for all 'TOR' server to switch links.
- Link lengths: ~50% > 3 m
  - From: *flatman_01_0911_NG100GOPTX.pdf*, reproduced with kind permission of Alan Flatman

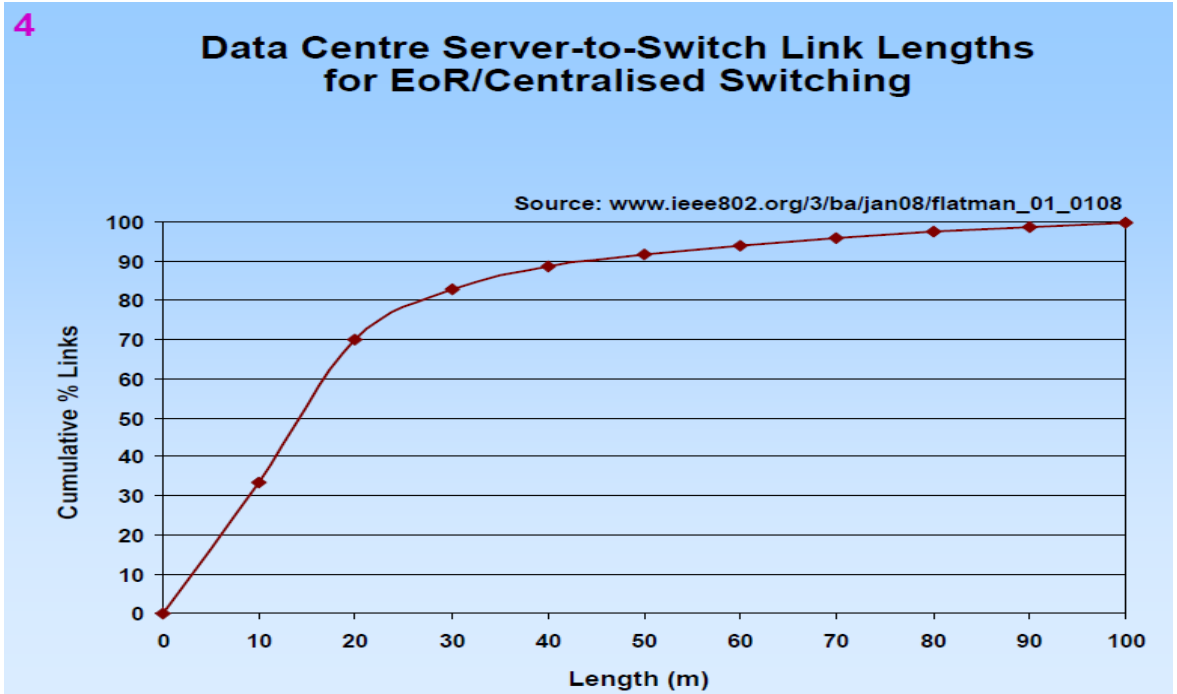


Consistent with the TIA-942 DC cabling standard limit (15m) for P2P cabling within the equipment distribution area (EDA) within and between racks.

http://www.ieee802.org/3/100GNGOPTX/public/sept11/flatman_01_0911_NG100GOPTX.pdf

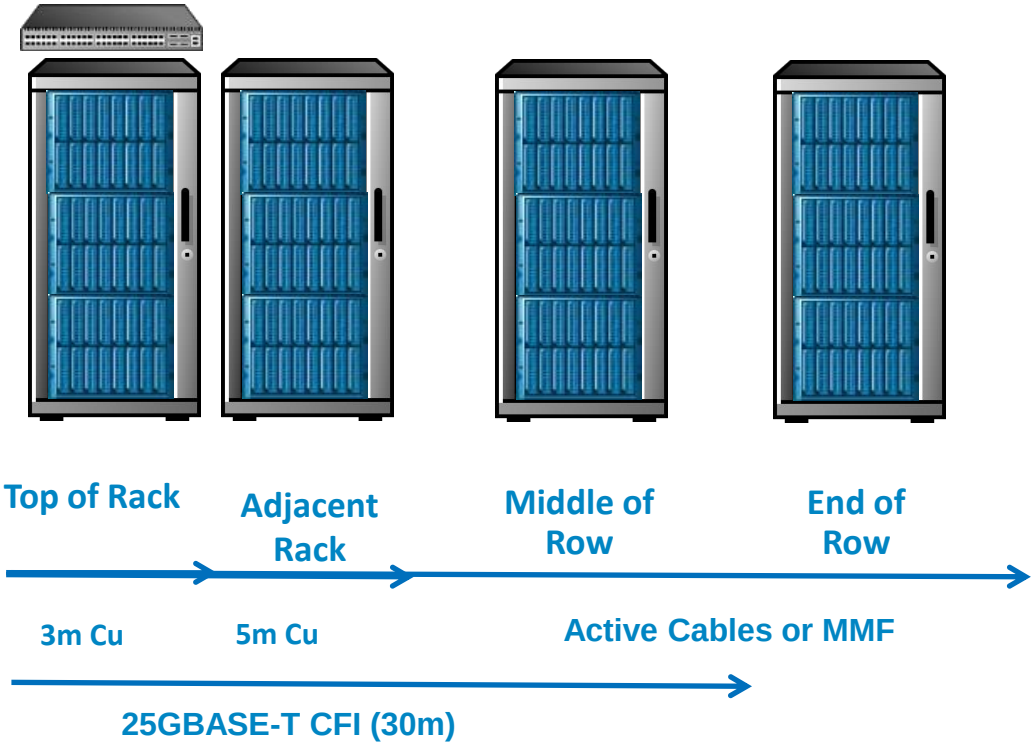
EoR link distributions

- Link lengths: ~90% > 3 m, ~83% < 30 m
  - From: *flatman_01_0911_NG100GOPTX.pdf*,



http://www.ieee802.org/3/100GNGOPTX/public/sept11/flatman_01_0911_NG100GOPTX.pdf

Example: What Servers Need from 25GbE

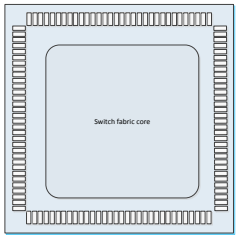


- No one solution for "server to switch" interconnect
  - Blade servers
  - Cu Cables
  - Active Cables
  - Multi-mode fiber
  - Future – 25GBASE-T?
- "Right" media dictated by customers architectural choices

Source: Gavin Cato, VP of Network Engineering, Dell, at Ethernet Alliance Technology Exploration Forum, October 16, 2014

25Gb/s I/O Efficiency

- Single 25 Gb/s lane per port connection from switch to PHY maximizes connectivity available in single ASIC
- Switch ASIC Connectivity limited by serdes I/O
- 25Gb/s lane maximizes bandwidth/pin and switch fabric capability vs. older generation
- 25Gb/s port optimizes both port count and total bandwidth for server interconnect



For a 128 lane switch:

Port Speed (Gbps)	Lane Speed	Lanes /port	Usable ports	Total BW (Gbps)
10	10	1	128	1280
25	25	1	128	3200
40	10	4	32	1280
40	20	2	64	2560
100	25	4	32	3200

Using 25Gb/s ports maximizes connectivity and bandwidth.

Market Summary

- The server interconnect world is very diverse.
  - ToR is right for some users, but for many applications and user facilities, EoR or MoR architectures enable better cost effectiveness, floor planning and support of mixed applications.

- Adding 25GBASE-T to the Ethernet Portfolio provides a transition roadmap from 1G->10G->25G->40G on compatible infrastructure
  - Majority of Server ports are still 1000BASE-T today
  - 10GBASE-T growing faster than twinax as Enterprise moves to 10G
  - Enables optimized switch port usage over broad range of server to switch architectures.
  - Enables large modular switches with high port counts.
  - Allows ports to be utilized which might otherwise be stranded with a 5m reach.
  - Allows 4x the switch port density (vs. 40GBASE-T) by supporting single lane MAC-to-PHY interface

25GBASE-T Technology Feasibility

George Zimmerman, CME Consulting

Technical Feasibility

- 25GBASE-T builds on well known technology deployed in the field for 10GBASE-T and being developed for 40GBASE-T
- 25GBASE-T could use the same 2-connector ISO Class I/Class II or TIA Category 8 channels currently defined in Draft 1.0 of IEEE P802.3bq
- Presentations in 802.3bq have shown technical feasibility at 40GBASE-T using 10GBASE-T based signaling at 3.2GHz symbol rate, one alternative is to run the same signaling at 2.0GHz symbol rate
- Technical approaches well understood from 10GBASE-T and 802.3bq:
  - Reuse of 10GBASE-T technology provides well-known models for feasible implementation
  - PHY channel models to 2GHz developed in 802.3bq allow estimation of PHY transmitter, receiver & cancellation parameters, SNR margins, including key blocks for managing power

Feasibility Comparison

Parameter	10GBASE-T (ref- Cl55)	40GBASE-T (ref- Cl 98d1.0)	25GBASE-T (example)
Channel	100m, Cl 55.7 (e.g., Cat6a)	30m, Cl 98.7d1.0 (e.g., Cat 8)	30m, Cl 98.7d1.0 (e.g., Cat 8)
Baud (MHz)	800	3200	2000 (example)
RX_ENOB (bits)	9.5-10	7.8	6.5-7.5
Channel IL (dB, Nyquist)	46.9	29.4	22.6
Channel round trip (baud)	880	1056	660
Echo Cancellation (dB)	55	47 (-6dB) to 55	43 (-12dB) to 55
NEXT Cancellation (dB)	40	34 (-6dB) to 40	28 (-12dB) to 40
FEXT Cancellation (dB)	25	22 (-3dB) to 25	19 (-6dB) to 25
Relative SNR Margins (dB)	0 (ref)	+2.7dB to 0.2 dB	+8.7dB to +4dB

Better than existing projects on most parameters

25GBASE-T – Why now?

Why Now?

- Enterprise Data Centers will begin transition to >10G speeds over the next few years
  - 25G and 40G will coexist
- Leverage investment in 10GBASE-T & 40GBASE-T technology
  - Provide roadmap from 10GBASE-T -> 25GBASE-T -> 40GBASE-T
- 25Gb/s technology provides optimal MAC-PHY interface
  - There are no 40Gb/s single lane standardization efforts under way… yet…
- The Ethernet Ecosystem has been very successful
  - Open and common specifications
  - Ensured Interoperability
  - Security of development investment

Proposed Path Forward

- Resources to develop 25GBASE-T reside in P802.3bq
- Propose extending the work of P802.3bq to include 25GBASE-T
  - Modest incremental effort, anticipate common technology
  - P802.3bq straw poll show unanimous support for that path
    - P802.3bq Sept'14 minutes

Proposed Next Steps

- 802.3 approve formation of 25GBASE-T Study Group (this plenary)
- SG determines objectives for 25GBASE-T
- SG prepares updated Objectives & CSD to modify the scope of P802.3bq
- 802.3 approval of modified P802.3bq project documentation (ideally March'15)

Questions?

Straw Polls

Call-for-Interest Consensus

- Should a study group be formed for 25GBASE-T?
- Y: 33+27=**60** N: **1** A: 3+5=**8**

- Room count:47+35=82

Participation

- I would participate in a 25GBASE-T study group in IEEE 802.3
  - Tally:18+19=37

- My company would support participation in a 25GBASE-T study group
  - Tally:13+12=25

Thank you!

Appendix: Diversity of DC architectures – references

from king_25GE_02a_0914.pdf

a. 'Navigating the Pros and Cons of Structured Cabling vs. Top of Rack in the Data Center', CCCA , May 2013, *http://www.cccassoc.org/files/2114/0138/9928/WP-SCS_vs_ToR_WP_Final_May_2013.pdf*

b. '40GBASE-T advantages and use cases', Andy Jimenez, 802.3bq, July 2014, jimenez_3bq_01_0711.pdf

c. '25GE server to switch architectures', Scott Kipp, 25G Ethernet Study Group Meeting, Ottawa, Sept 2014, kipp_25GE_01_0914.pdf

d. 'Deploying 10GBASE-T as the low cost path to the cloud', Panduit, Intel & Cisco joint white paper, *http://www.panduit.com/ccurl/413/996/zcan02_10gbase_tecosystem.pdf*

e. 'Reconsidering Physical Topologies with 10GBASE-T', Broadcom white paper, May 2013, *http://www.broadcom.com/collateral/wp/84848-WP100-R.pdf*

f. 'Link distance and server connectivity', Scott Kipp, Next Generation Base T study group, Nov 2012, *http://www.ieee802.org/3/NGBASET/public/nov12/kipp_01a_1112_ngbt.pdf*

MoR & EoR topologies allow fewer stranded ports on servers and on switches.
Substantial cost-effectiveness improvements.
Better flexibility for thermal planning in MoR and EoR architectures.
Support of mixed applications.
Support of reconfiguration.