

Use Case Analysis for 40GBASE-T

Bob Wagner – Panduit

IEEE P802.3bq 40GBASE-T Task Force

Beijing, China

March, 2014

Introduction

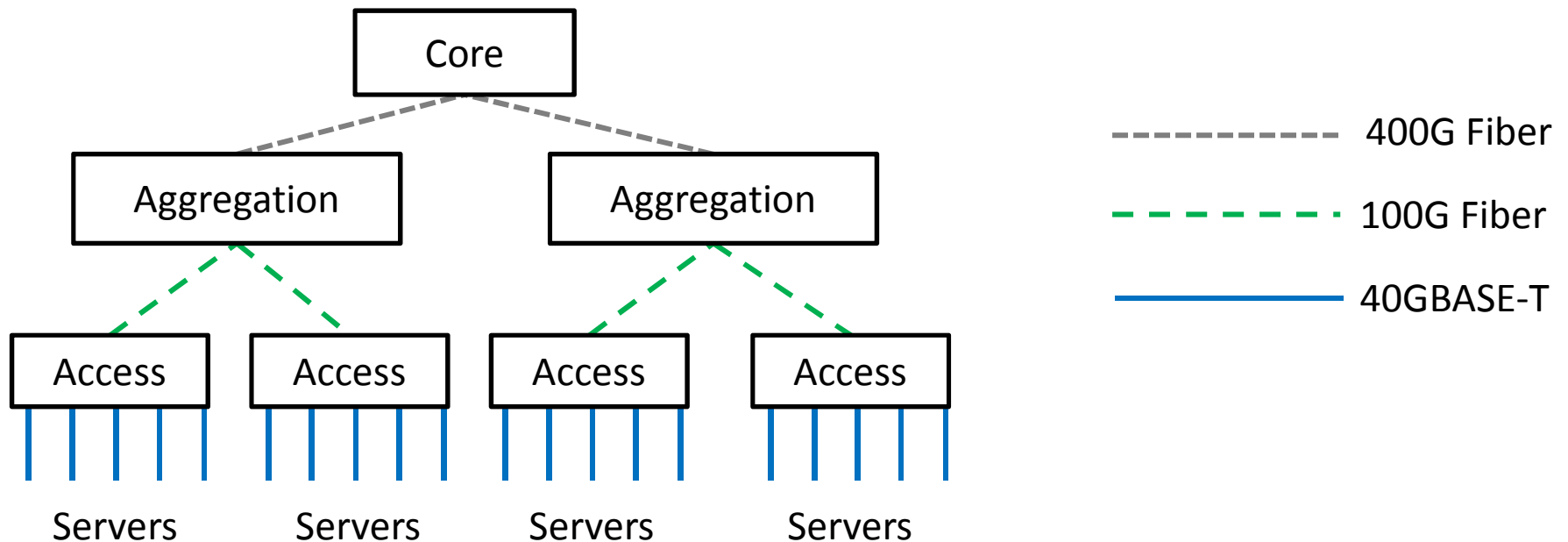
- BASE-T Ethernet widely recognized as offering the benefits we all know
 - Lowest cost
 - No transceivers
 - High volume easy to manufacture
 - Easy installation
 - Field termination
 - Auto-negotiation (backwards compatibility, plug and play)
 - Sufficient reach
 - 100m not always needed
- Does 40GBASE-T offer these same advantages with the 30m limitation?

Analysis

- Understand architectures used today and those intended for tomorrow
 - How is copper used in these architectures?
- For a given architecture:
 - Will 40GBASE-T work with the length limitation?
 - Understand relative cost of different designs for that architecture

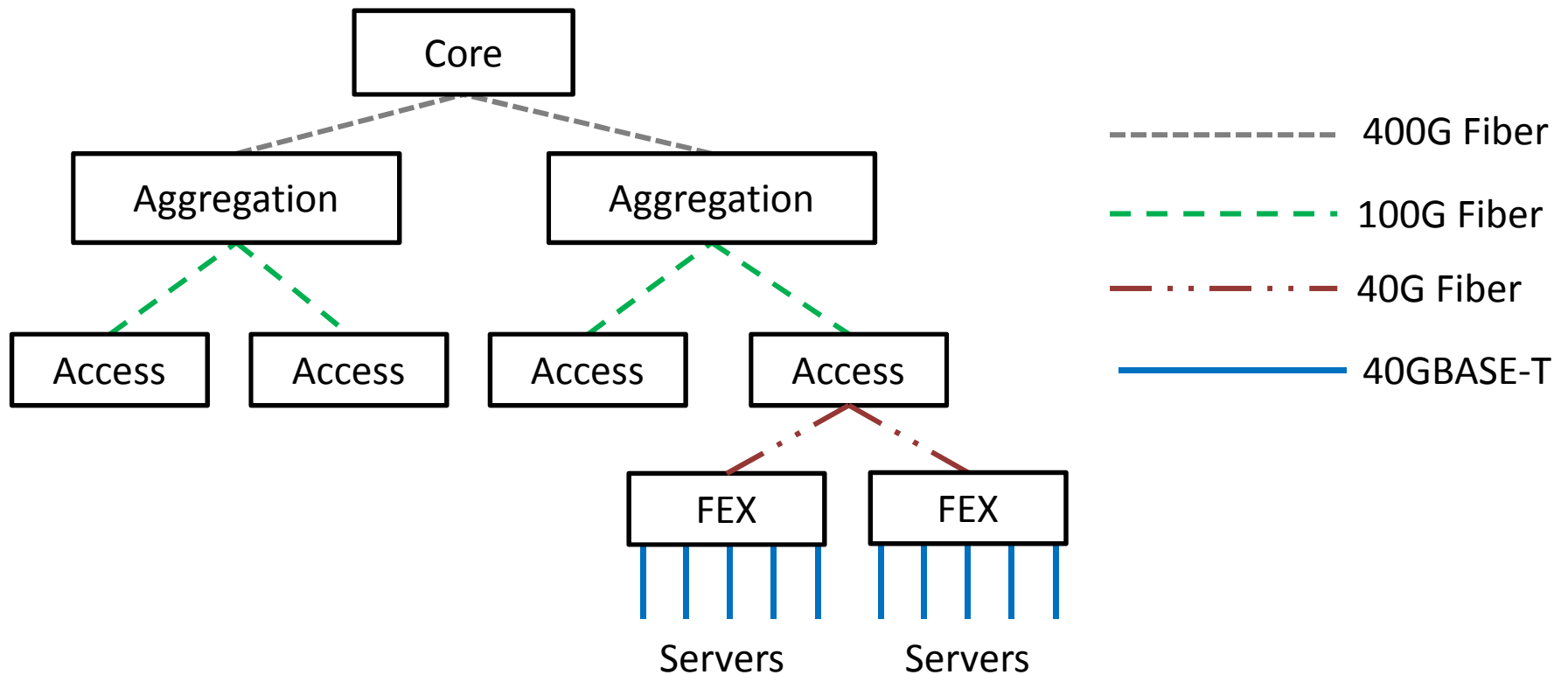
40G to Server Tiered Architecture

- 40GBASE-T Used as Server Link
 - Supports ToR, EoR, or Network Rows for copper in data center physical designs
- Fiber used as Switch-to-Switch
 - Fiber required due to port options from equipment vendors

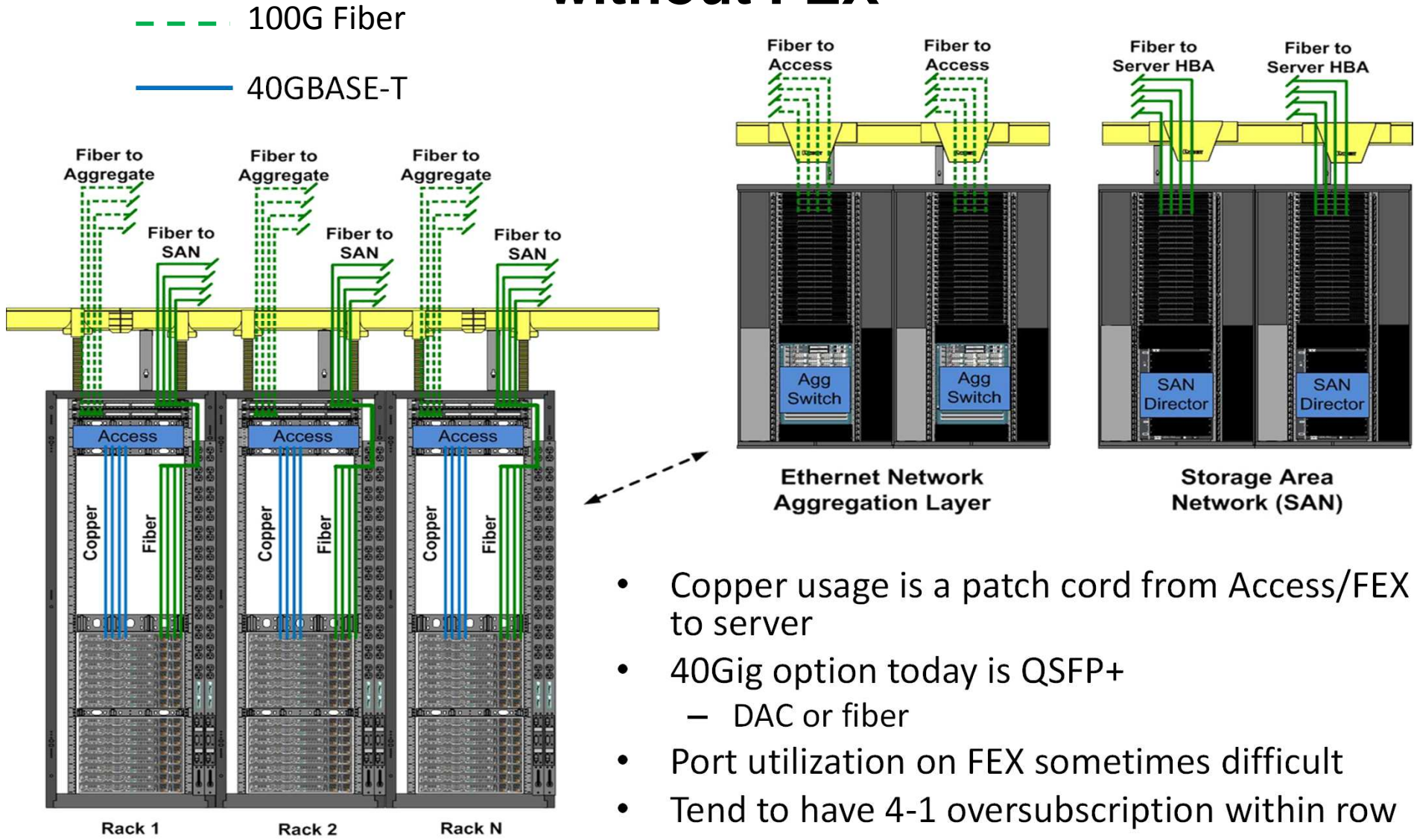


40G to Server Tiered Architecture with Fabric Extender (FEX)

- Copper Used as Server Link
 - Almost exclusively ToR with copper used only as a point-to-point patch cord
- Fiber used as Switch-to-Switch
 - Fiber required due to port options from equipment vendors

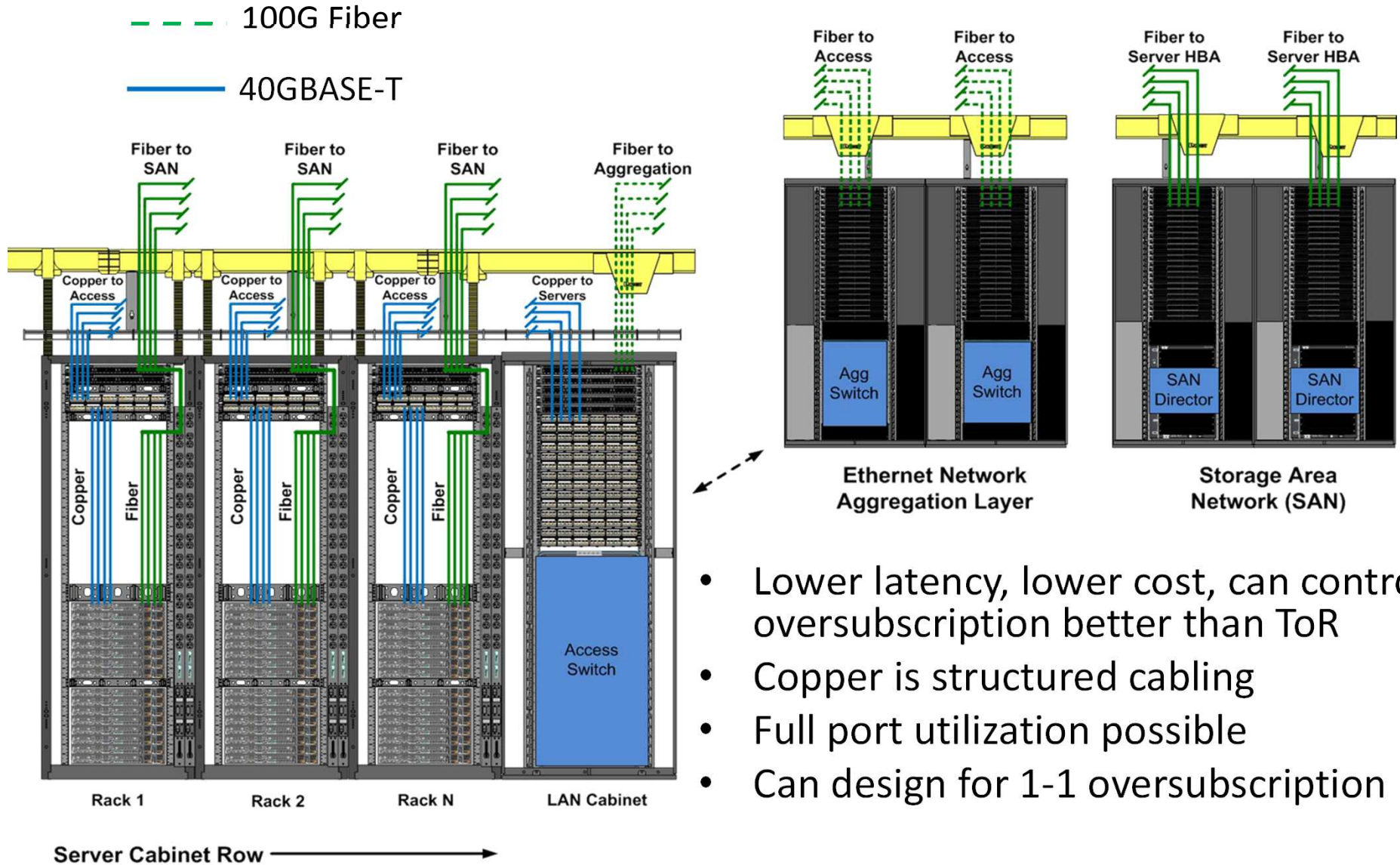


40G to Server ToR Data Center Design with or without FEX



- Copper usage is a patch cord from Access/FEX to server
- 40Gig option today is QSFP+
 - DAC or fiber
- Port utilization on FEX sometimes difficult
- Tend to have 4-1 oversubscription within row

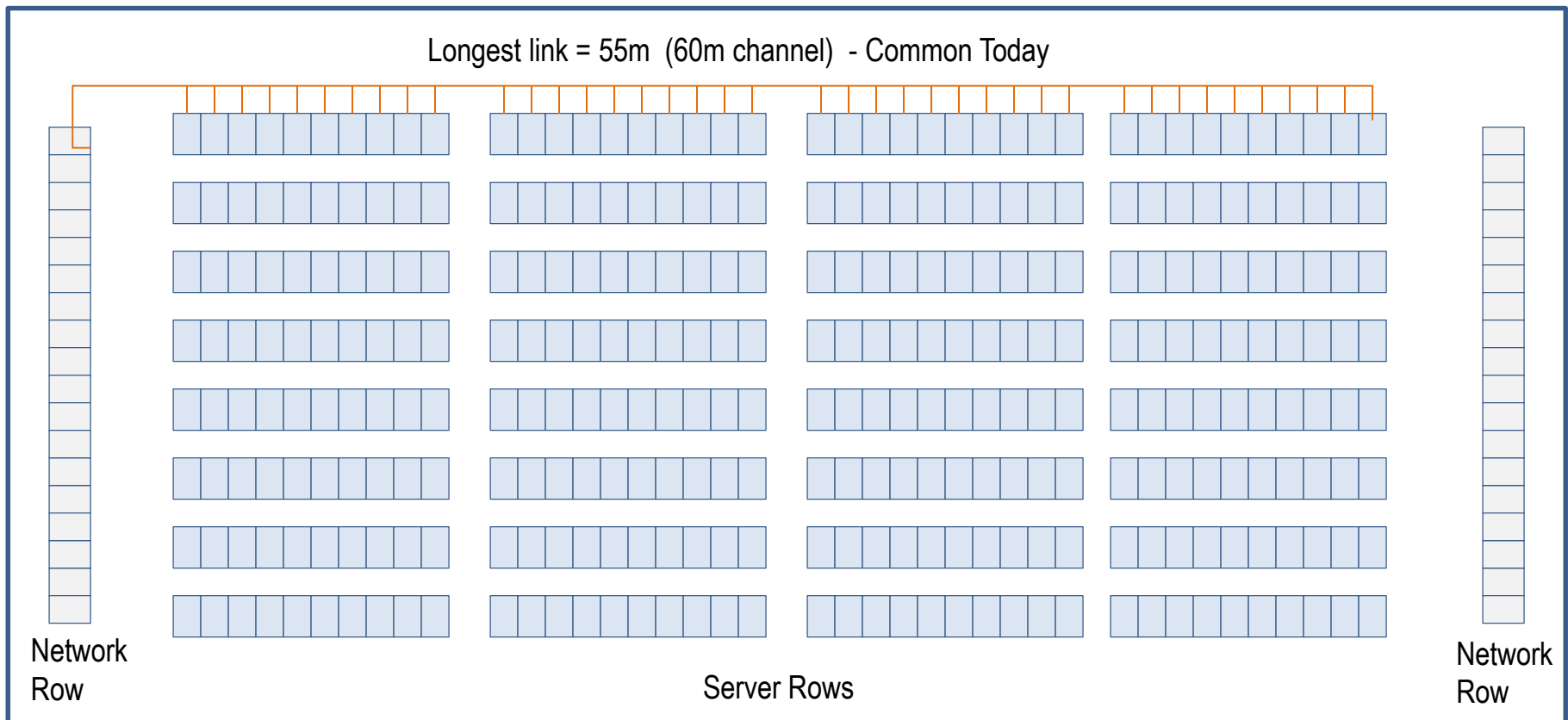
40Gig to Server EoR Data Center Design



- Lower latency, lower cost, can control oversubscription better than ToR
- Copper is structured cabling
- Full port utilization possible
- Can design for 1-1 oversubscription

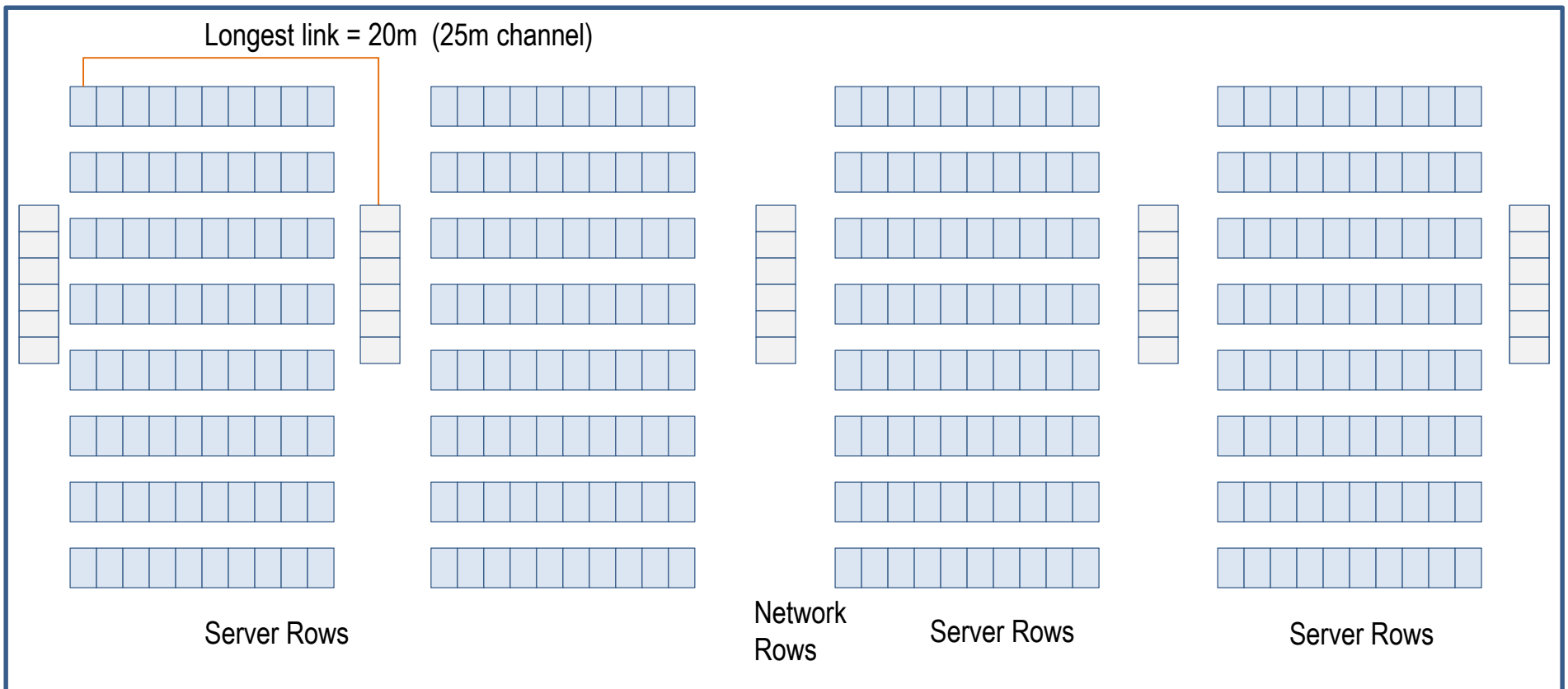
Network Row

- Copper used to connect switch to server
- Copper length needed determined by number of servers



Network Row Redesigned

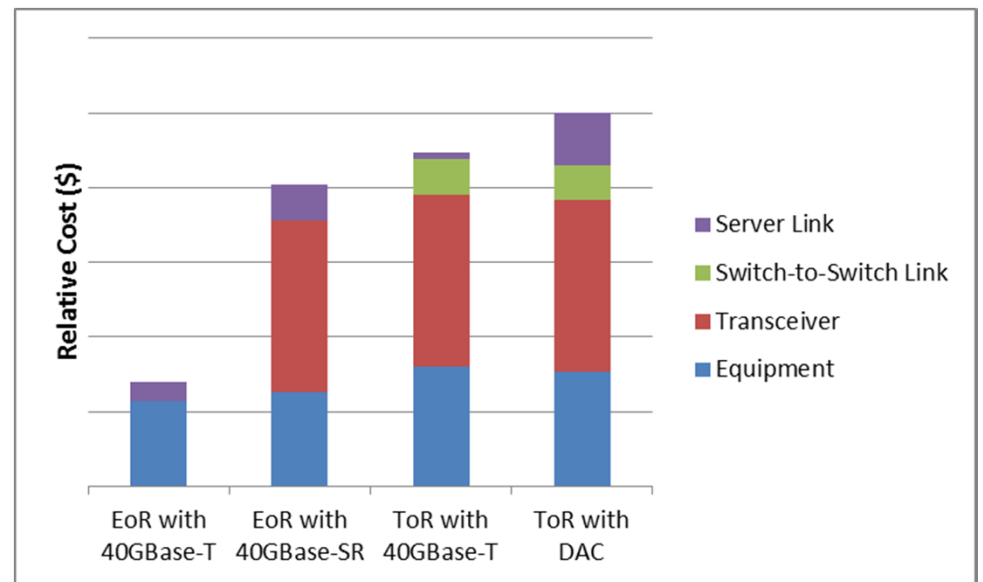
- Rather than having 2 at each end, have in middle
- Having a 30m reach is only limiting in brownfields where infrastructure needs 100m



Network Design Comparisons

40GBASE-T Use Case	Needed Reach	Sufficient?	Relative Cost
Top of Rack	~5m	Yes	Medium
End of Row	~30m	Yes	Low
Network Row	~70m	Yes, with optimized data center layout	Low

- Looking at relative cost per port from access layer down
- Note EoR has 1-1 oversubscription, ToR has 4-1
- Transceivers drive cost



Summary for Tiered

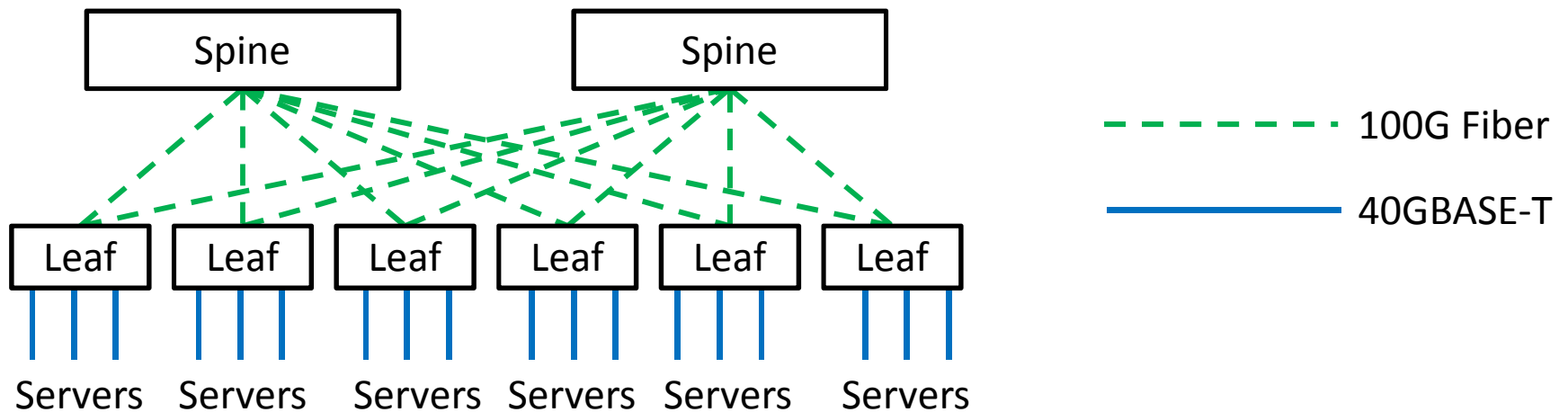
- 40GBASE-T retains same advantages as traditional Base-T
 - Same benefits in cost
 - Same benefits in easy of use
- Any perceived reach limitations can be overcome on greenfield installations with new optimized layout design
- 40GBASE-T will work with traditional 1000BASE-T and 10GBASE-T advantages
- What about new applications?
 - Assume uplinks at 100G

Leaf-Spine

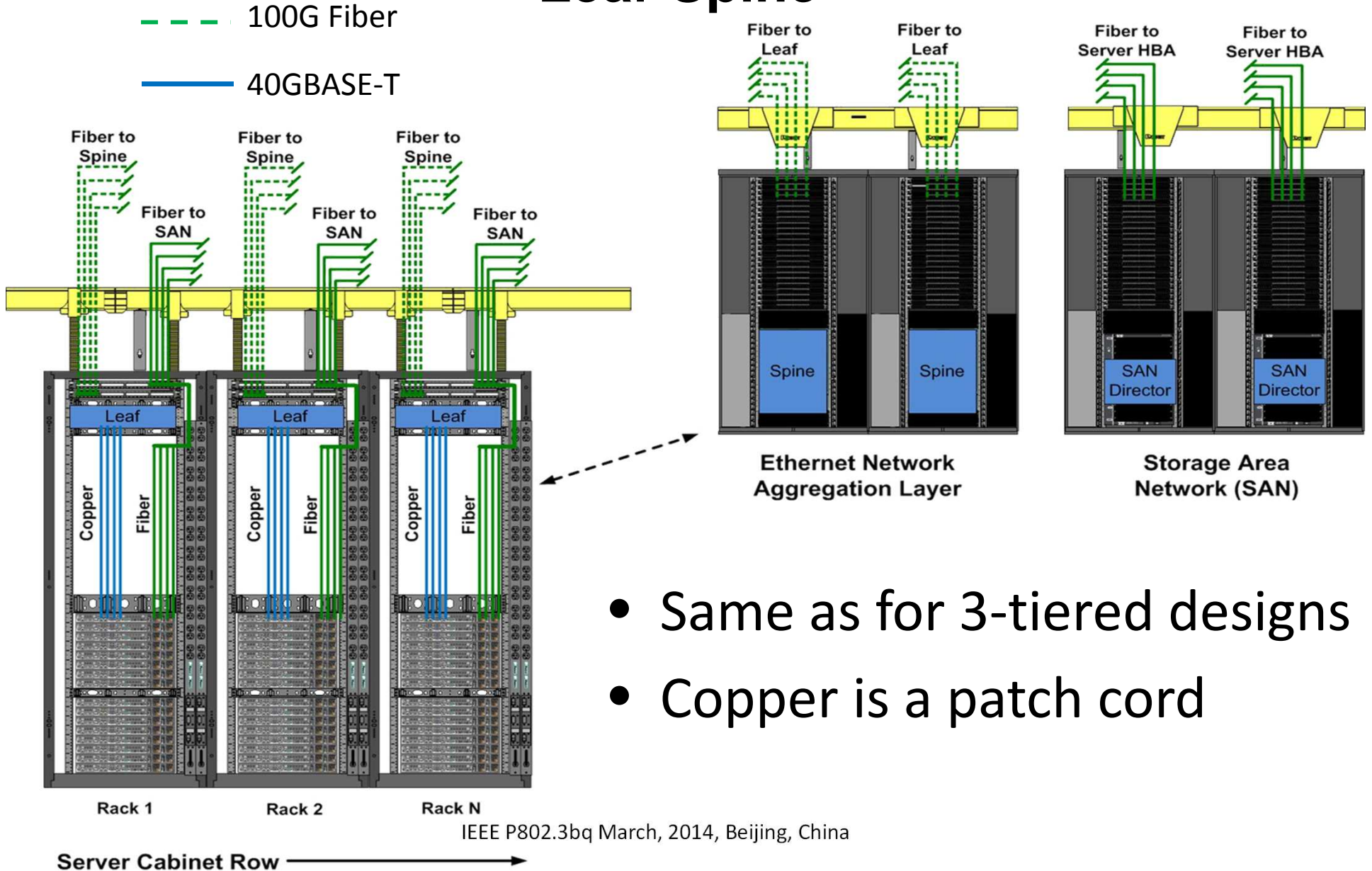
- Leaf-spine is the new flatter architecture for data centers
 - Only 2 layers – leaf and spine
- How does 40GBASE-T fit into that?
- Does it provide any unique advantages?

40G to Server Leaf-Spine Architecture

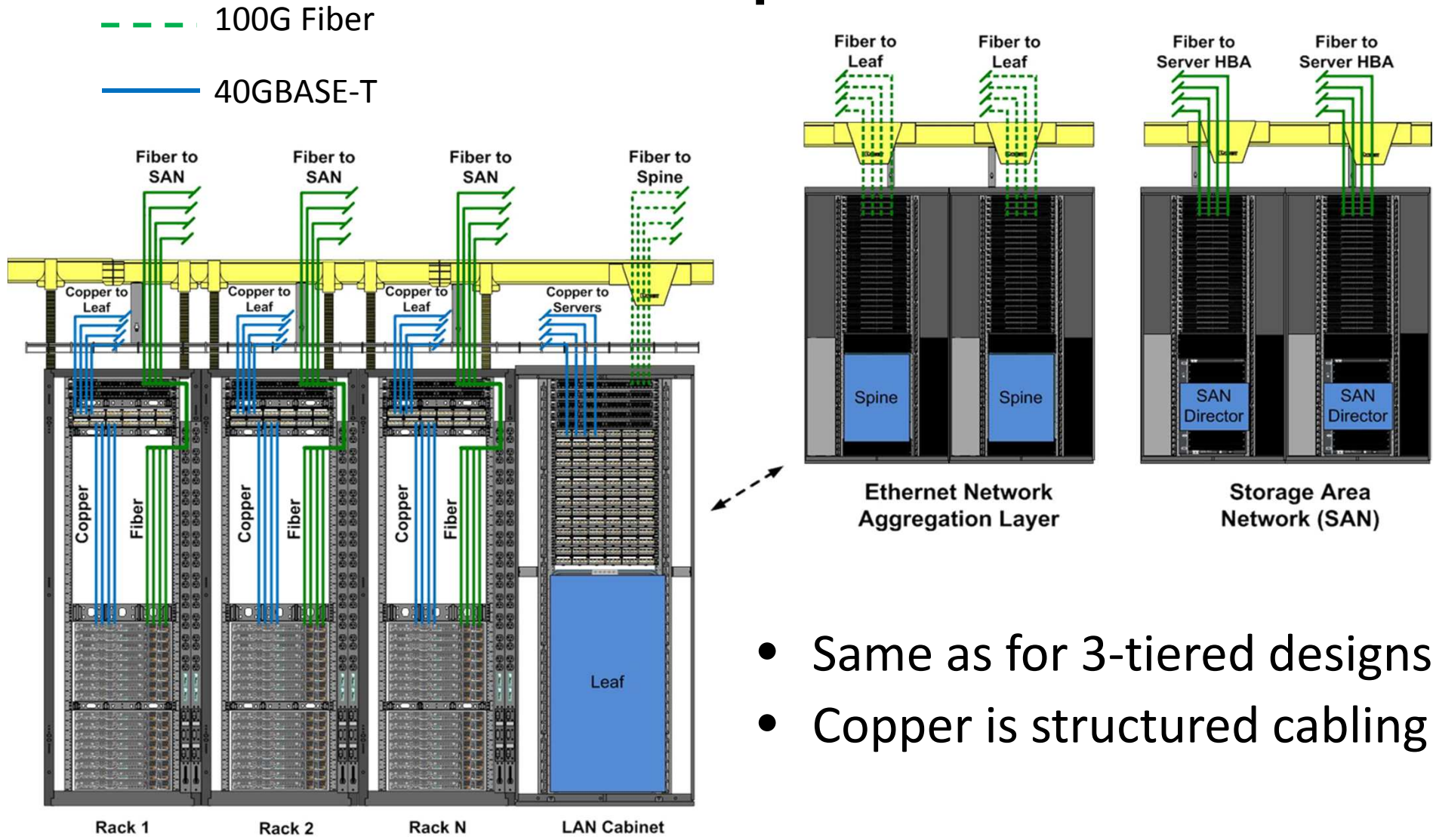
- Copper still used as server link
 - Intended to be used in ToR or EoR
 - Copper used because it is the cost effective option
- Fiber used as Switch-to-Switch
 - Today most uplinks (40G) require QSFP modules (BiDi or MPO/MTP)



40G to Server ToR Data Center Design with Leaf-Spine



40G to Server EoR Data Center Design with Leaf-Spine



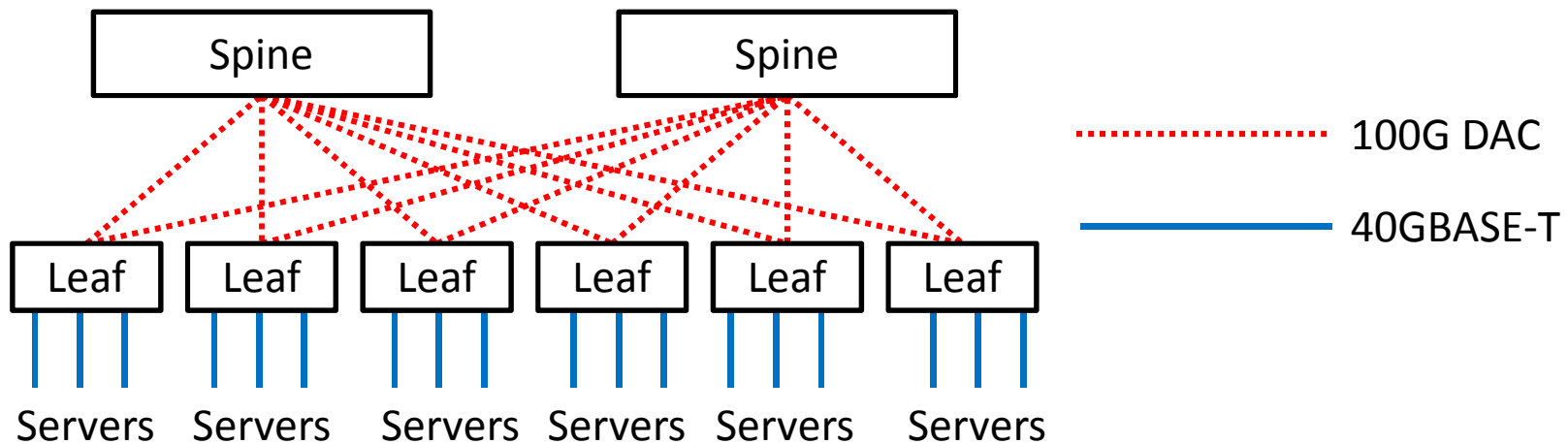
- Same as for 3-tiered designs
- Copper is structured cabling

Leaf-Spine

- Leaf-Spine 40GBASE-T offers essentially the same cost advantages when using copper as for 3-tiered designs
- However, the leaf-spine design presents opportunities

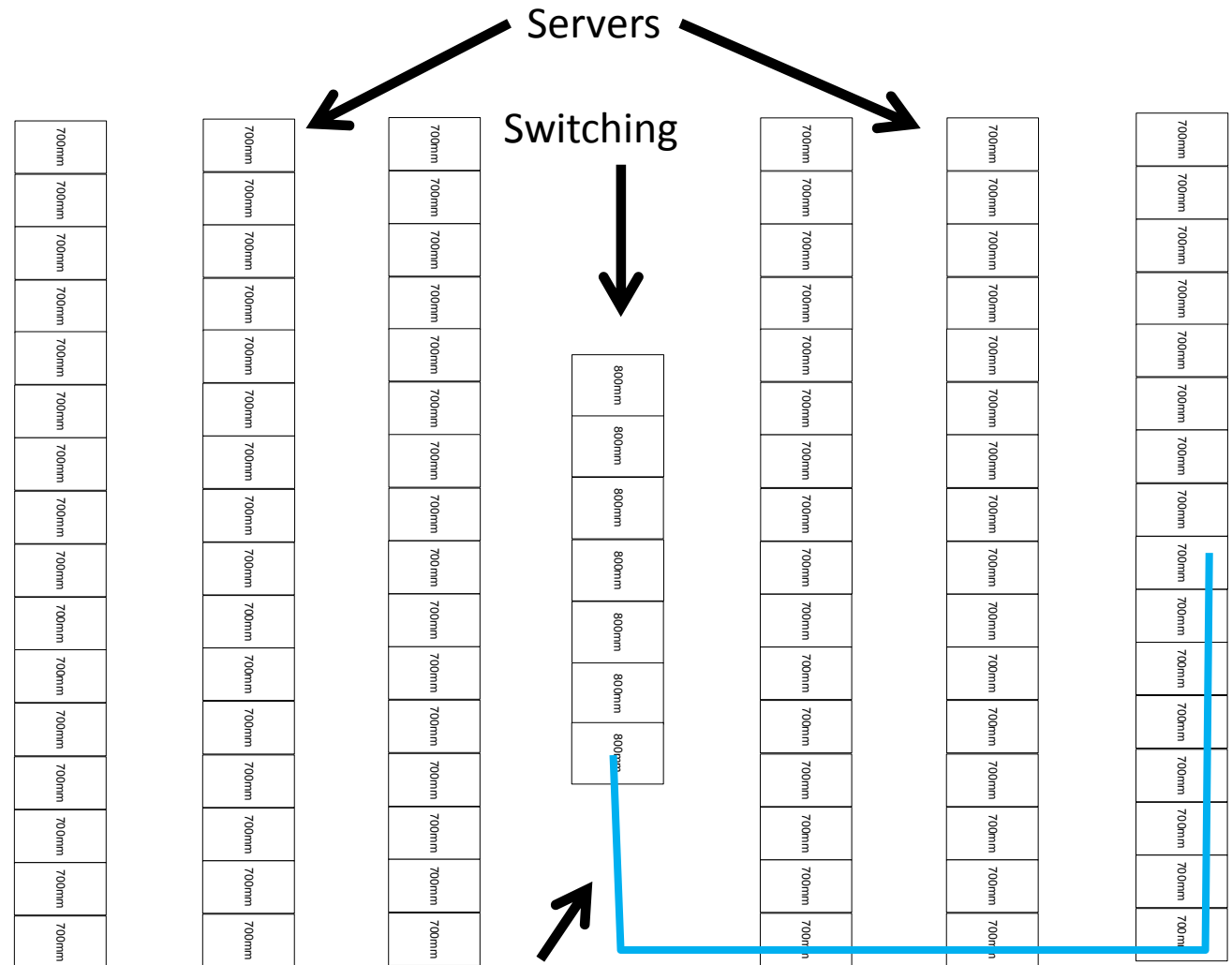
40G to Server Leaf-Spine with Centralized Switching

- Copper is the server link using structured cabling
- Direct attach copper (DAC) or fiber (AOC) used to connect servers that are located together
 - Removes cost of transceivers



Leaf-Spine with Centralized Switching, Hot/Cold Aisle – 2304 servers in 30m reach

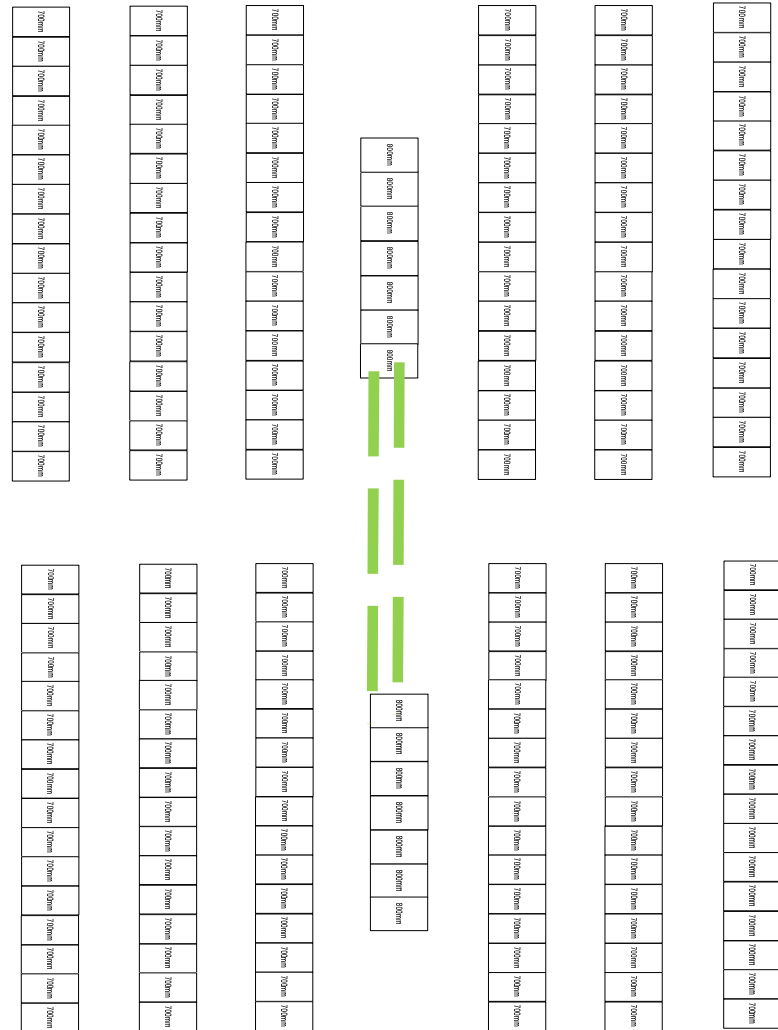
- 30m reach limits to 2304 servers
 - 24 servers / cabinet
 - 16 cabinets / row
 - 6 rows
- Preserves hot / cold aisle



Worst-case cable routing within 30m reach limitation

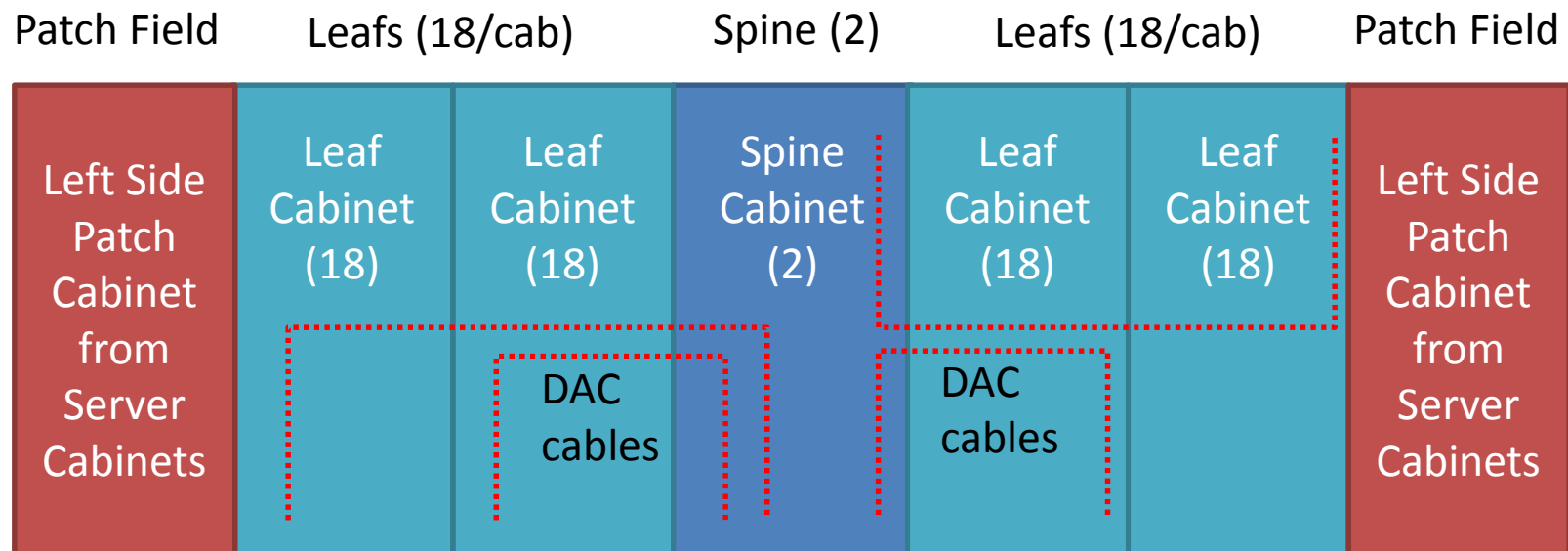
Scalability of Leaf-Spine with Centralized Switching

- Can add several more 2304 server pods by linking Network rows together



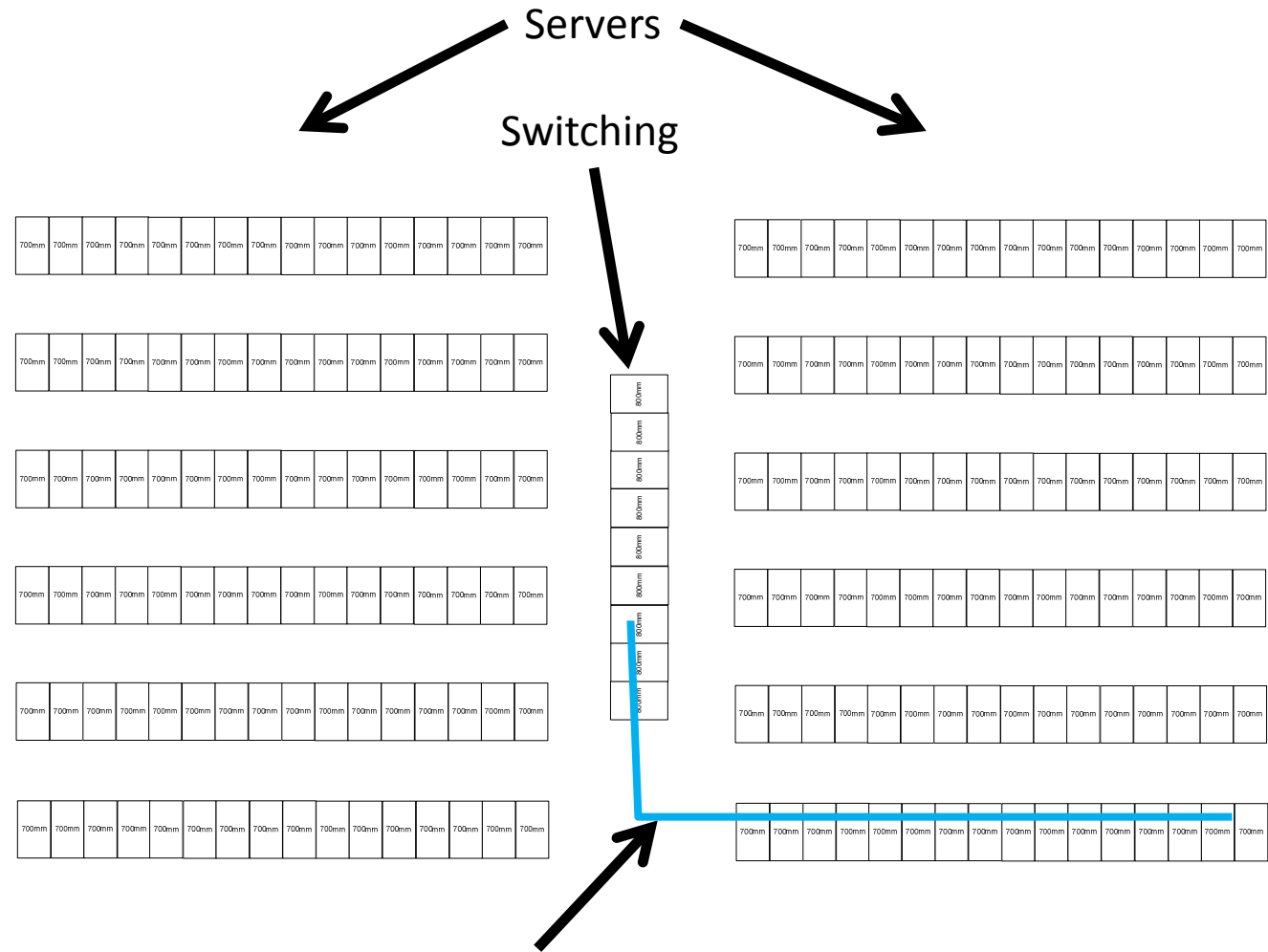
5m DAC allows for 3456 servers

- All leaf-spine connections within 5m DAC reach
 - Connections occur in middle of cabinet with cabinet sidewalls removed
- 2 spines, 72 leafs (18/cabinet, 4 cabinets), 48 40GBASE-T ports / leaf
 - 3456 servers possible



Alternate Concept – 3456 servers w/ DAC

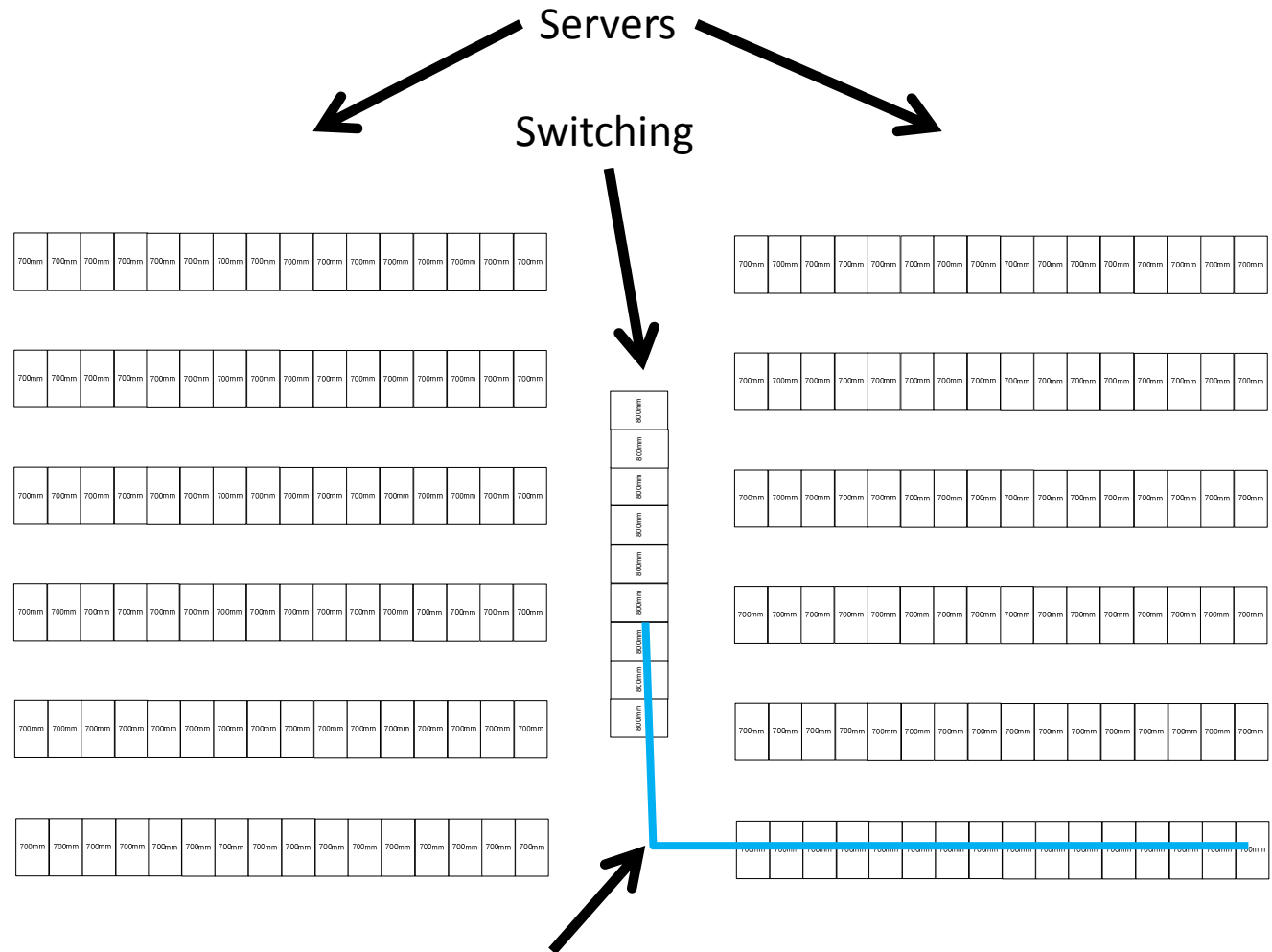
- 30m reach limits to 4608 servers
 - 24 servers / cabinet
 - 12 cabinets / row
 - 12 rows
- Cold air flooded DC with chimneys



Worst-case cable routing within 30m reach limitation

Alternate Concept – 4608 servers using AOC

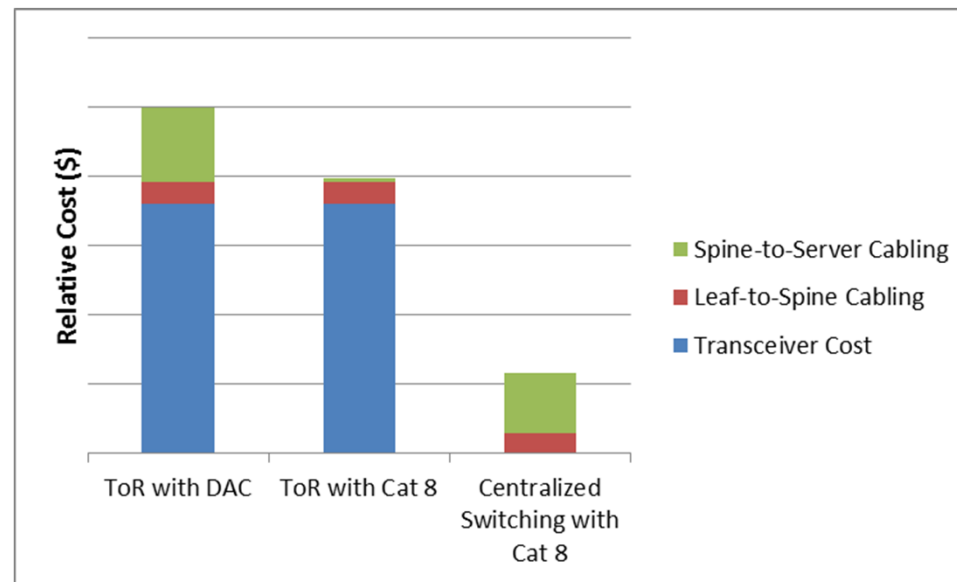
- 30m reach limits to 4608 servers
 - 24 servers / cabinet
 - 16 cabinets / row
 - 12 rows
- Cold air flooded DC with chimneys



Worst-case cable routing within 30m reach limitation

Centralized Switching Costs

- Comparing the cost of the connectivity and cabling in 3 options
 - Includes both server link and switch to switch link
- Compares ToR with DAC, ToR with 40GBASE-T, and Centralized Switching
- Relative costs looking at 3456 servers



Centralized Switching Summary

- Data centers can be designed for this TODAY and future proofed for TOMORROW
- Presents large cost savings for small to mid-sized data centers
- Will drive adoption of new leaf-spine technology with lower cost
- Methods to expand size of centralized switching model

Questions or Comments?