BROCADE

LINK DISTANCE AND SERVER CONNECTIVITY FOR 20M 100GBE OPTICS

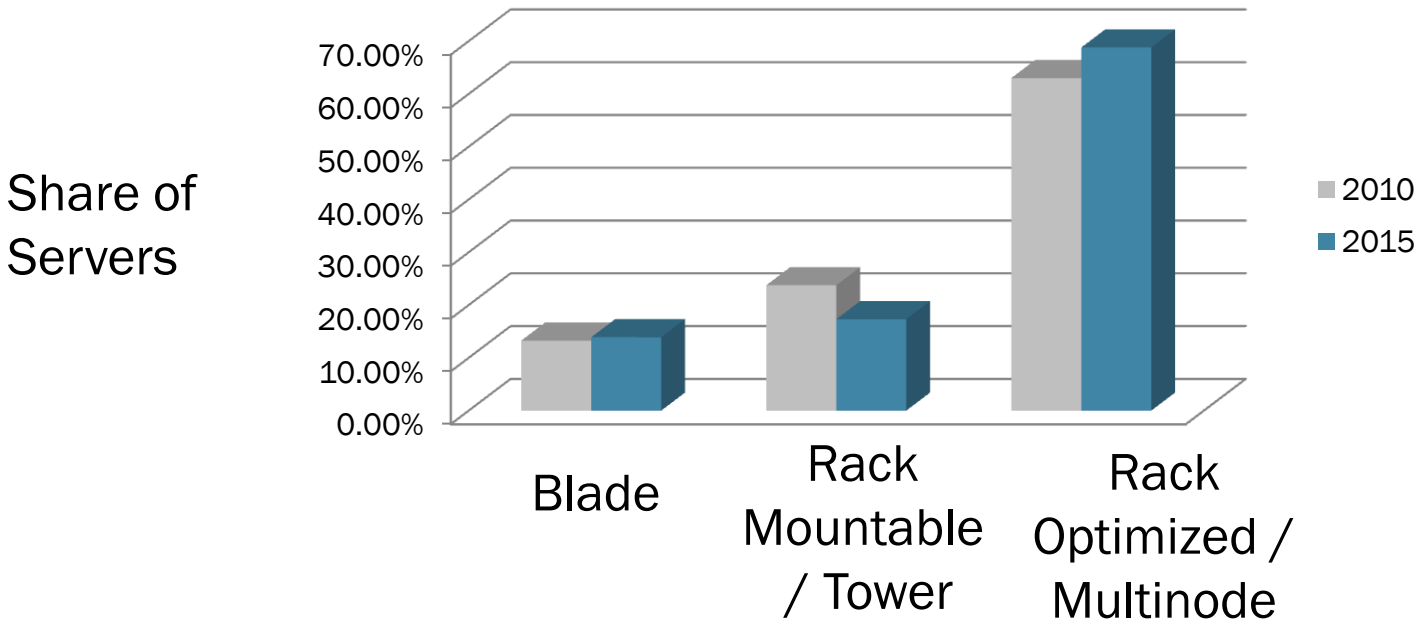
Scott Kipp

skipp@brocade.com

November 2012

Trends in Servers

- The 20 m objective was set mainly for supercomputing applications
- Most servers being sold today are rack optimized
- Rack optimized servers are very dense and their connectivity can be supported with relatively short links



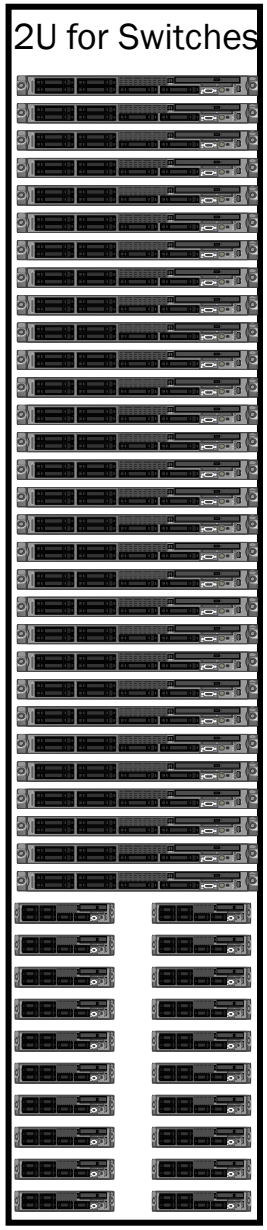
Source: Gartner Forecast: Server by Form Factor, Worldwide 3Q12 Update

How Many servers are in a rack?

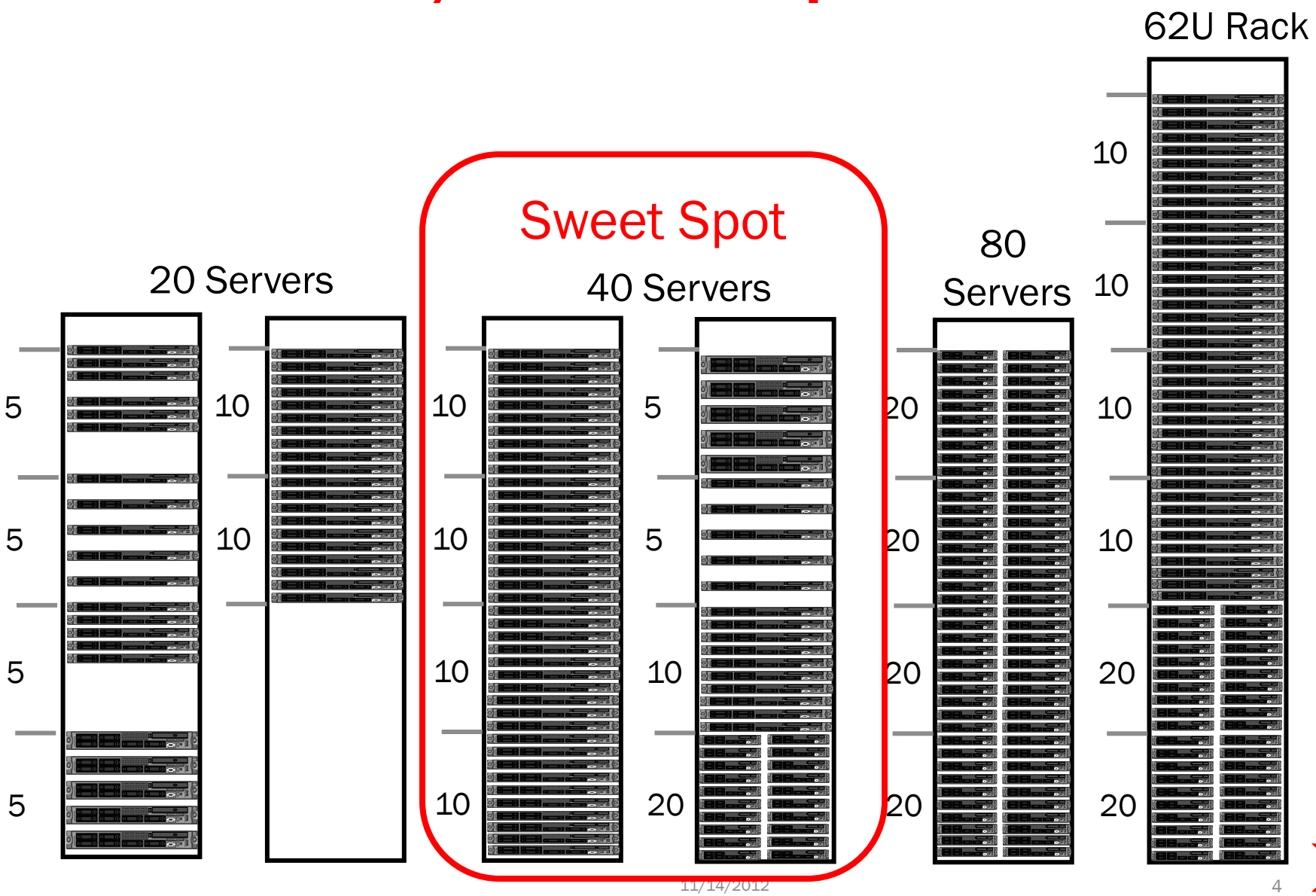
- 42U racks are the most common racks and can support:
  - 20 2U servers and 2 ToR (Top of Rack) switches
  - 20 1U servers with other equipment or limitations in the rack
  - 40 1U servers and 2 ToR switches
  - 80 1/2U servers and 2 ToR switches – multi-node servers, 64-bit ARM servers and other low power servers are coming in less than 1U

40 Servers /Rack

80 Servers /Rack



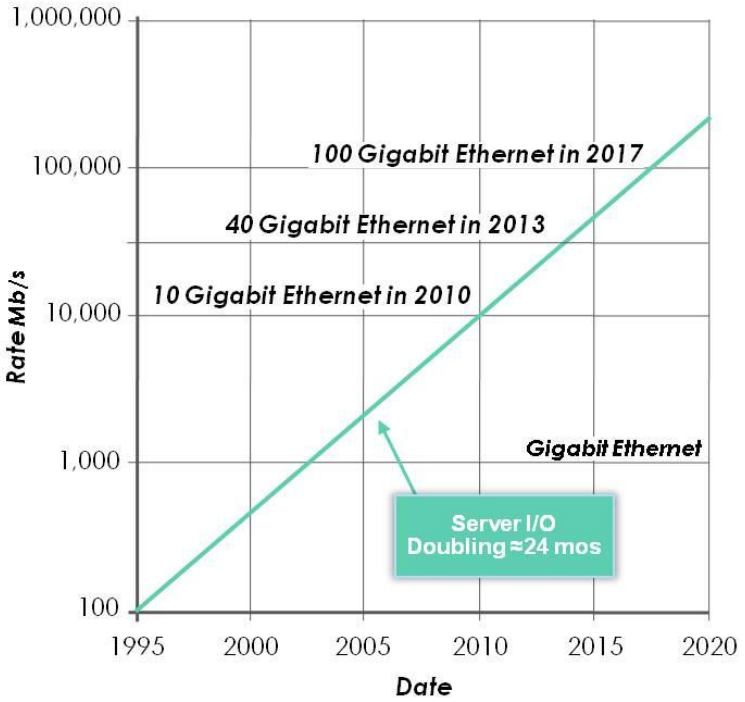
20-80 Servers / Rack Examples



Bandwidth Within Rack

- Rack throughput easily reaches Terabits/second

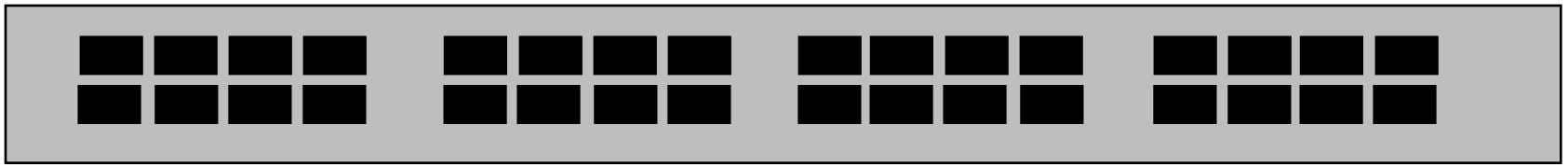
	Bandwidth / Rack (Gbps)		
Bandwidth / Server	20 Servers / Rack	40 Servers / Rack	80 Servers / Rack
10G/server	200	400	800
20G/server	400	800	1600
40G/server	800	1600	3200
80G/server	1600	3200	6400
100G/server	2000	4000	8000



Source: Figure 1 of Bandwidth Assessment Report

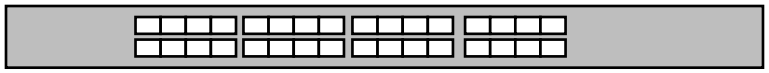
ToR Switch Bandwidth

- Image a 1U ToR switch has 32 100GBASE-SR4 Lite ports (3.2Tbps), then it might have 12 uplinks for 20 downlinks to servers
- You would need one to four of these switches per rack
- Higher 100G port counts are more than 5 years out



32 100G QSFP28 or CFP4
Ports = 3,200 Gpbs
32 ports X 3W/Port = 96W

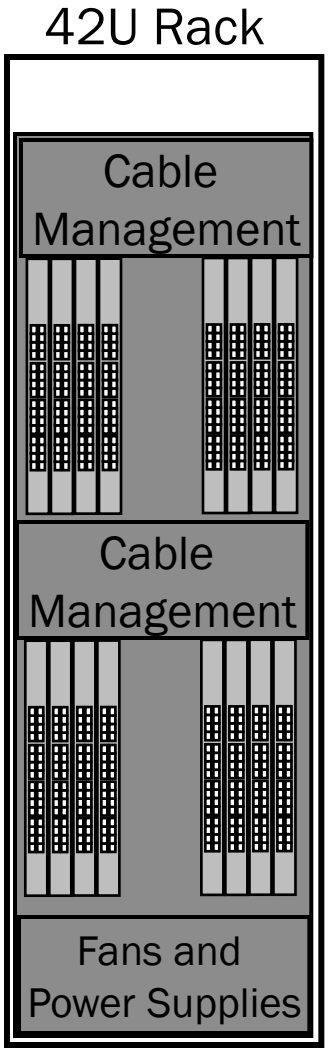
Modular Switch Bandwidth



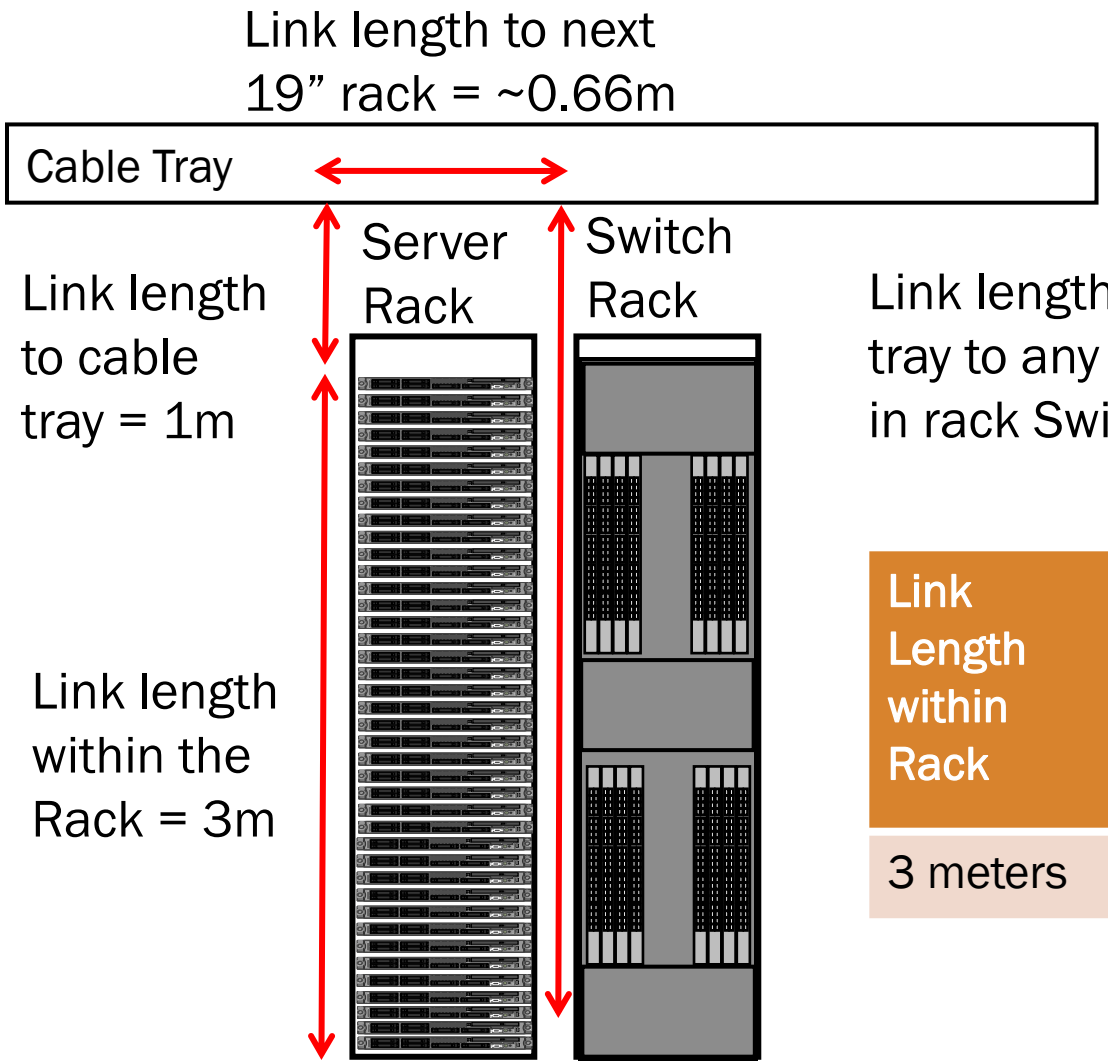
Blade for Modular Chassis with 32 100GBASE-R Ports = 3.2 Tbps

~16 blades typical in 40U chassis
12 blades / chassis for 32 ports of 100GBASE-R ports and 4 blades for uplinks
= 384 100GBASE-R ports

40U Modular Chassis with 16 blades
~ 384 ports of 100GBASE-R Ports



Rack to Rack Link Distances

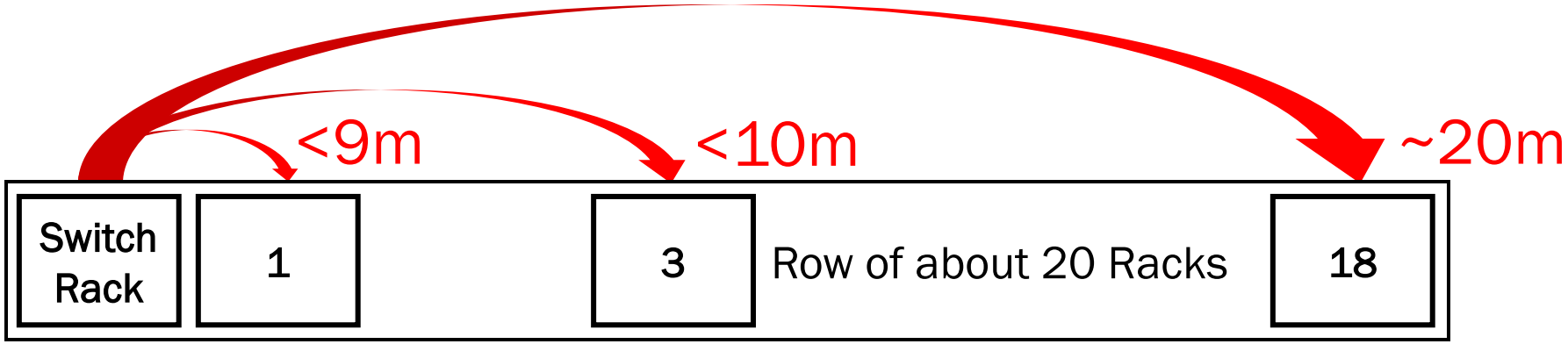


Link length from cable tray to any switch port in rack Switch = 4m

Link Length within Rack	Link length to Two Racks	For Each Additional Rack
3 meters	8.66 meters	+ 0.66m

Server Connectivity with End of Row Topology

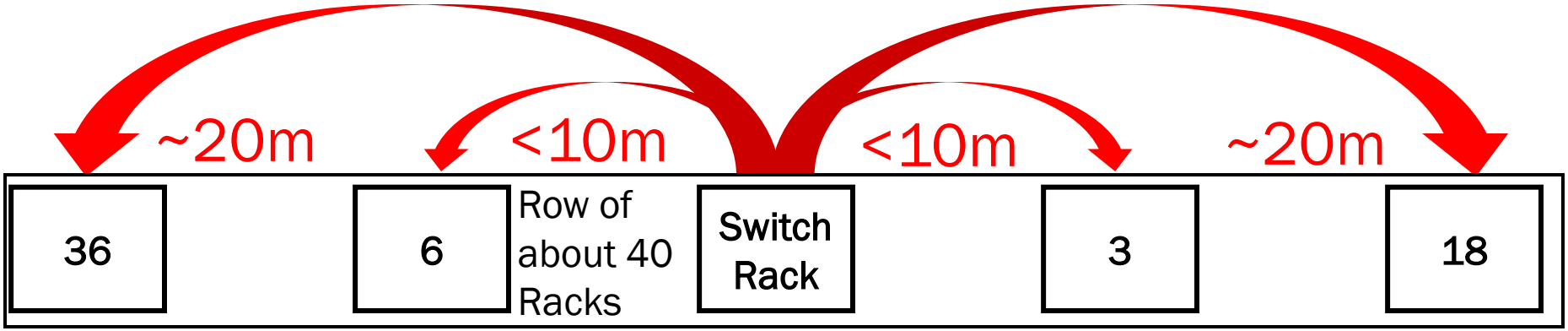
Reach 100s and ~1,000 servers with 20 meters



	1 Racks - 8.66m	3 Racks 10m	...	18 Racks 20 meters
20 Servers /Rack	20 servers	60 servers		360 servers
40 Servers /Rack	40 Servers	120 Servers		720 servers
80 Servers /Rack	80 Servers	240 Servers		1440 servers

Server Connectivity with Middle of Row Topology

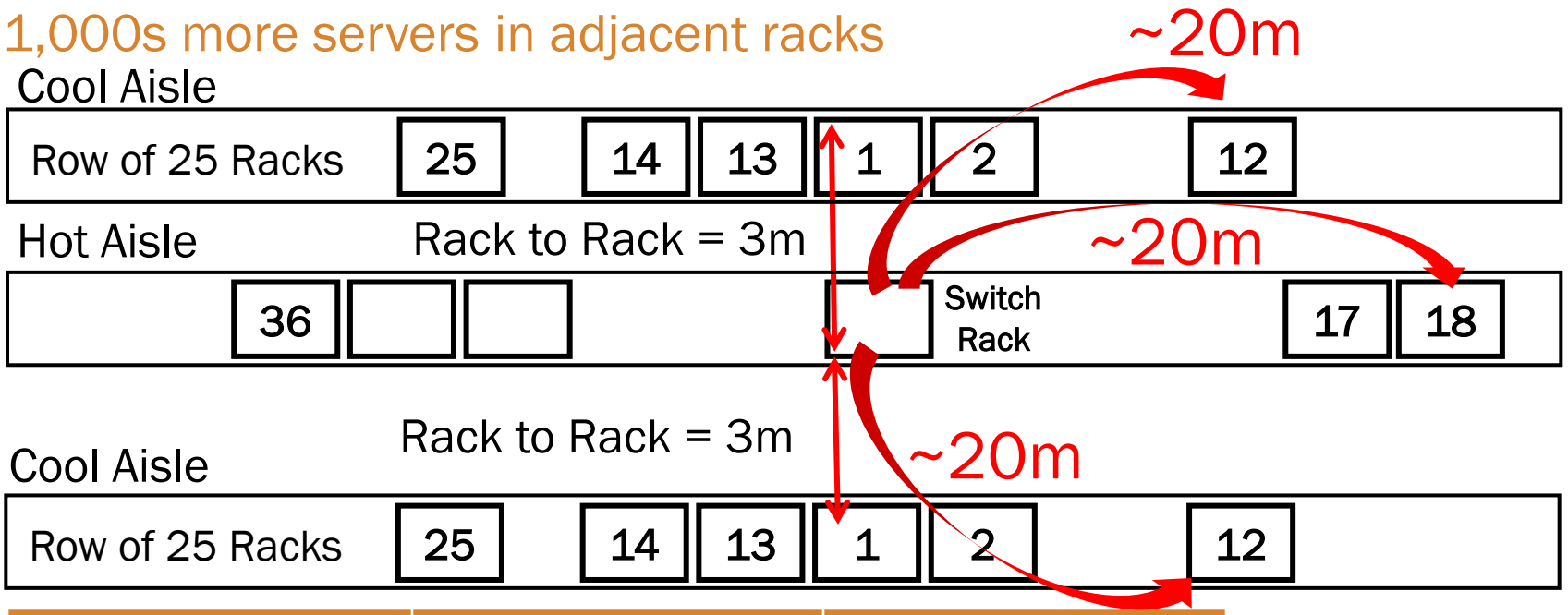
Reach 100s of servers with 10 meters, 1000s with 20m



	6 Racks - 10m	...	36 Racks 20 meters
20 Servers /Rack	160 servers		720 servers
40 Servers /Rack	320 Servers		1440 servers
80 Servers /Rack	640 Servers		2880 servers

Adjacent Aisle Server Support

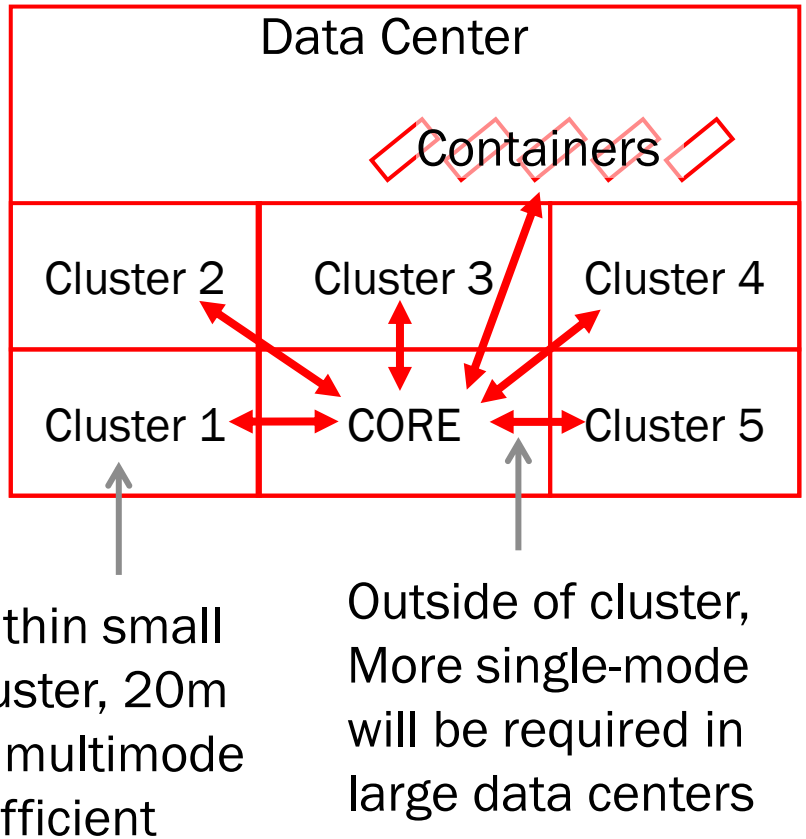
1,000s more servers in adjacent racks



	Adjacent Hot Aisle with 25 Racks	Adjacent Cool Aisle with 25 Racks
20 Servers /Rack	500 more Servers	+500 servers
40 Servers /Rack	1,000 more Servers	+1,000 servers
80 Servers /Rack	2,000 more Servers	+2,000 servers

Clusters and Data Centers

- A cluster is usually less than a few thousand servers that could be serviced with a 20 meter link and 1U servers
- Inter-cluster or data center links are growing in distance – more single-mode



Why not AOCs?

- Is structured cabling required in supercomputing?
- If not, can the loss budget be shrunken to almost 0db?
  - Insertion loss = 0 connector loss + 0.020km * 0.5dB/km = 0.01db
  - Does this make it lower cost than a 100 meter solution?
- If no structured cabling required, why not use Active Optical Cables (AOCs) for this reach?

BROCADE

THANK YOU