

# **LINK DISTANCE AND SERVER CONNECTIVITY**

Scott Kipp

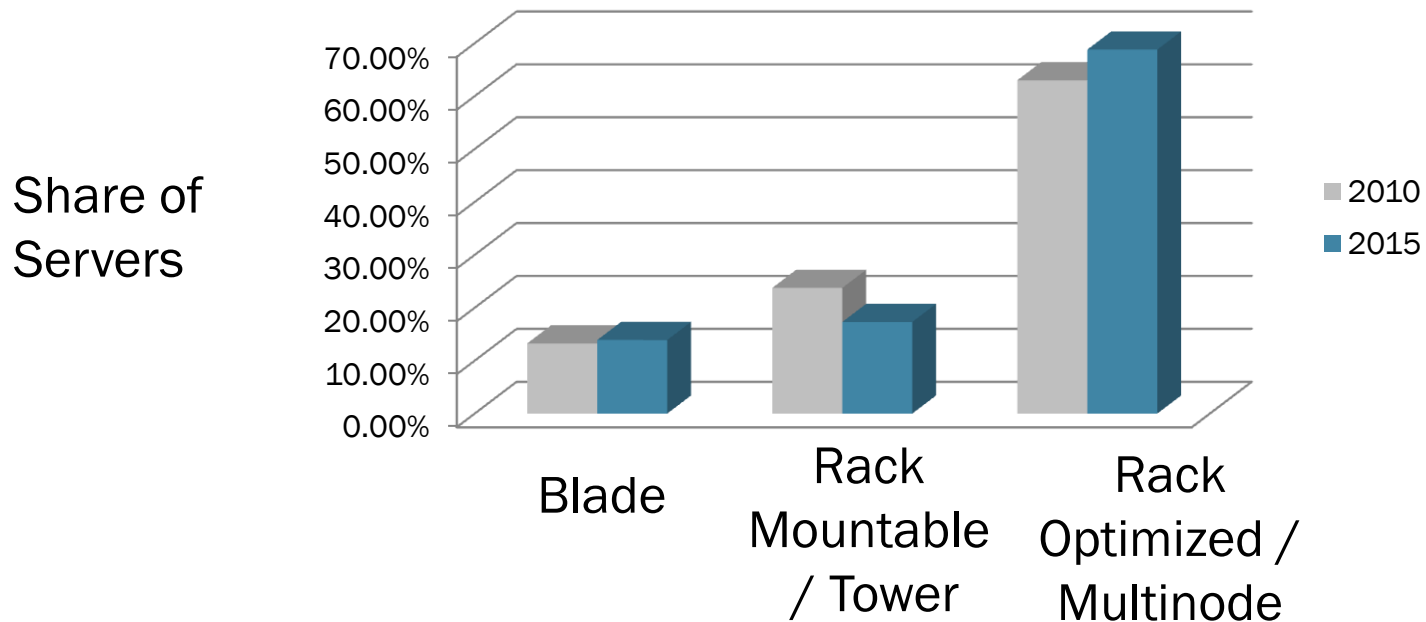
[skipp@brocade.com](mailto:skipp@brocade.com)

November 2012



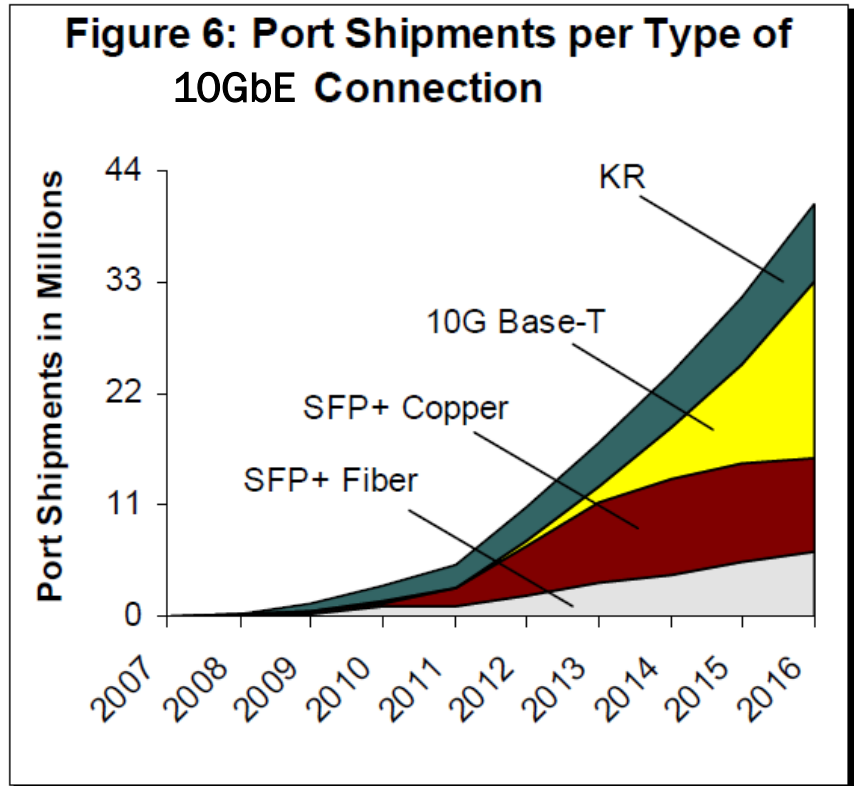
# Trends in Servers

- Most servers being sold today are rack optimized
- Rack optimized servers are very dense and their connectivity can be supported with relatively short links



# BASE-T is Dominant Server Interconnect

- 1000BASE-T is dominant server interconnect today
- 10GBASE-T is miniscule today, but growing fast
- 10GBASE-T has not been widely adopted 6 years after standardization because of high power consumption
- 40GBASE-T should not make the same choice



Source: Dell'Oro Controller and Adapter Forecast, January 2012



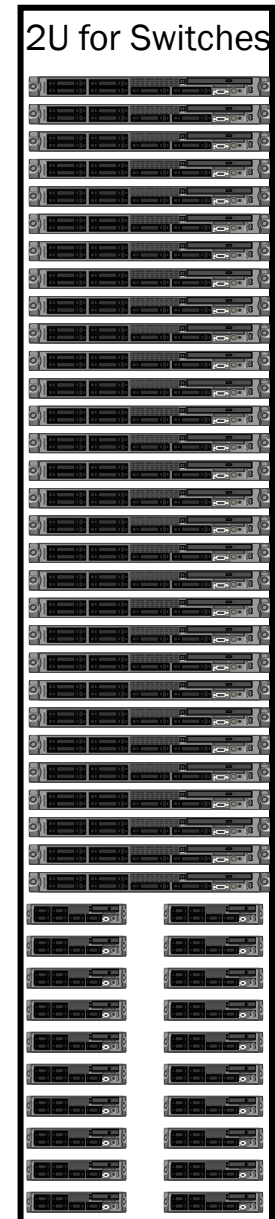
# How Many servers are in a rack?

- 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**



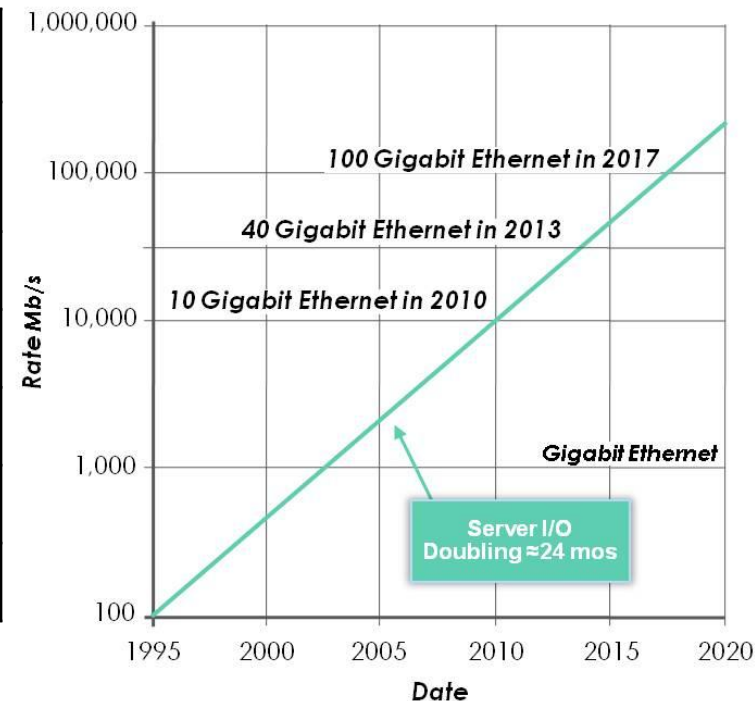
# 20-80 Servers / Rack Examples



# Bandwidth Within Rack

- Rack throughput easily reaches Terabits/second

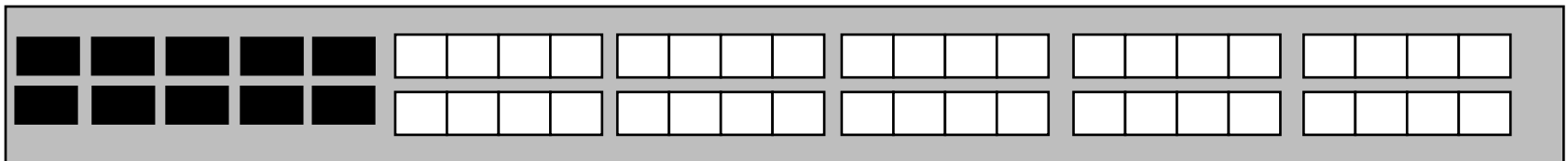
	Bandwidth / Rack (Gbps)		
Bandwidth / Server	20 Servers / Rack	40 Servers / Rack	80 Servers / Rack
10G/server	200	400	800
20G/server	400	800	1600
40G/server	800	1600	3200
80G/server	1600	3200	6400
100G/server	2000	4000	8000



Source: Figure 1 of Bandwidth Assessment Report

# ToR Switch Bandwidth

- If a 1U ToR switch has 40 40GBASE-T ports (1.6Tbps), then it might support up to a Terabit of uplinks
- 1U switch should stay under 500 Watts, and are usually under 250 W
- Without low power, 40GBASE-T will not reach this type of density:



10 100G QSFP28 or CFP4

Ports = 1,000 Gpbs

10 ports X 3W/Port = 30W

40 40GBASE-T Ports = 1,600 Gpbs

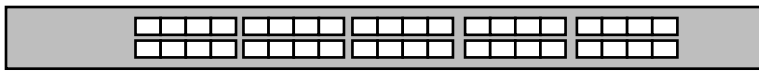
40 ports X 3W/Port = 120W

40 ports X 2W/ port = 80W

**150W dedicated to PHY is high power!**

**2W/40GBASE-T would drop this to 110W**

# Modular Switch Bandwidth



Blade for Modular Chassis with 40  
40GBASE-T Ports = 1.6 Tbps

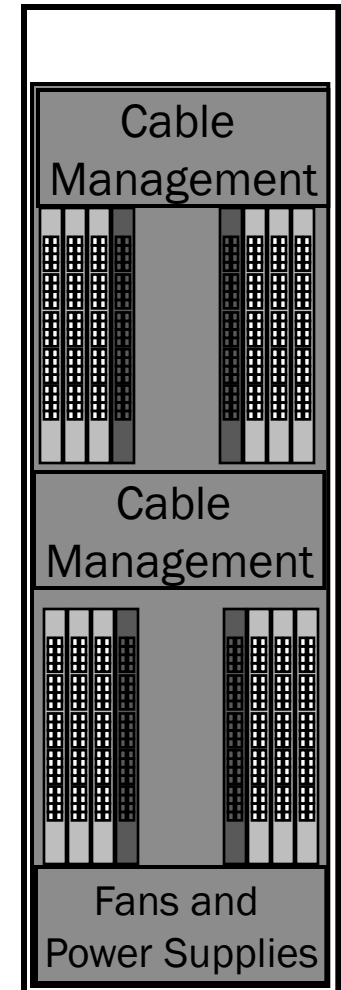
~16 blades typical in 40U chassis  
12 blades / chassis for 40 ports of  
40GBASE-T ports and 4 blades for  
uplinks

= 480 40GBASE-T ports

Max of 560 40GBASE-T ports with  
no uplinks

**40U Modular  
Chassis with 12  
blades of 40  
ports of  
40GBASE-T  
~ 480 ports**

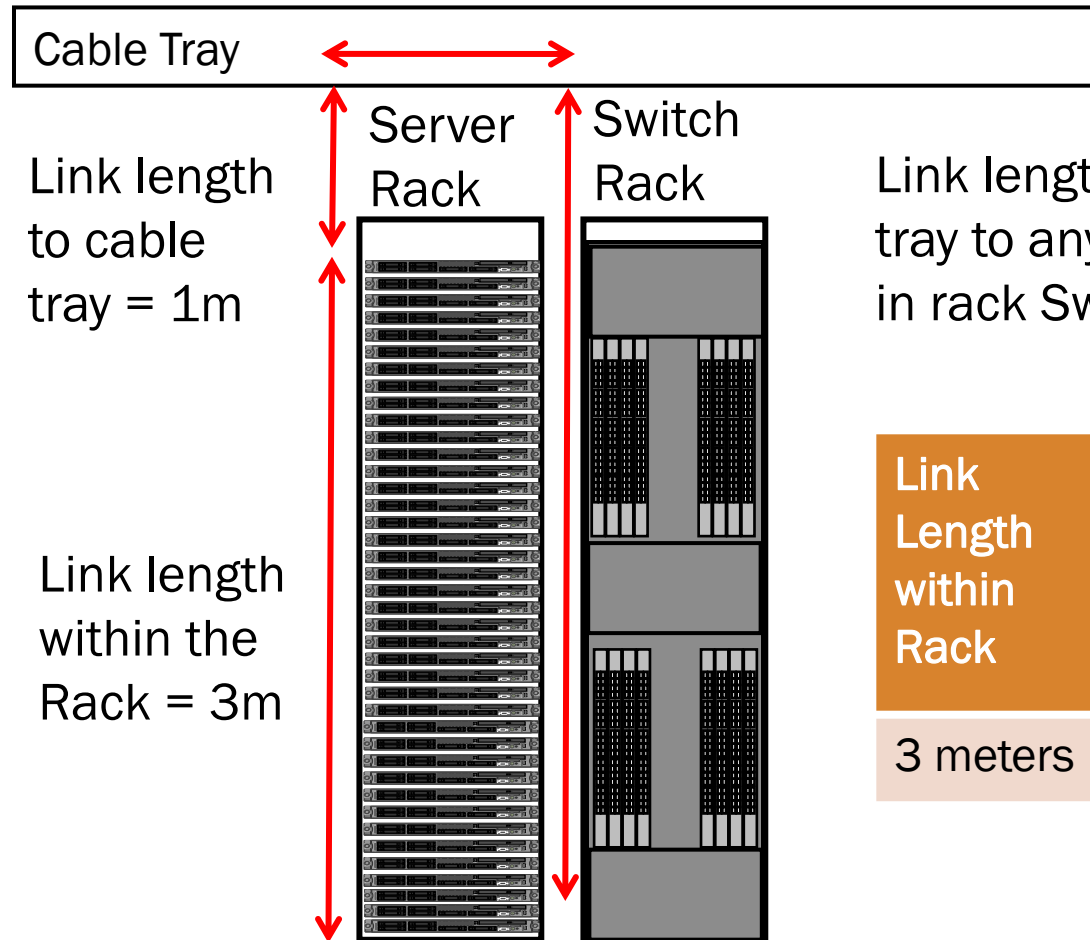
42U Rack





# Rack to Rack Link Distances

Link length to next  
19" rack = ~0.66m

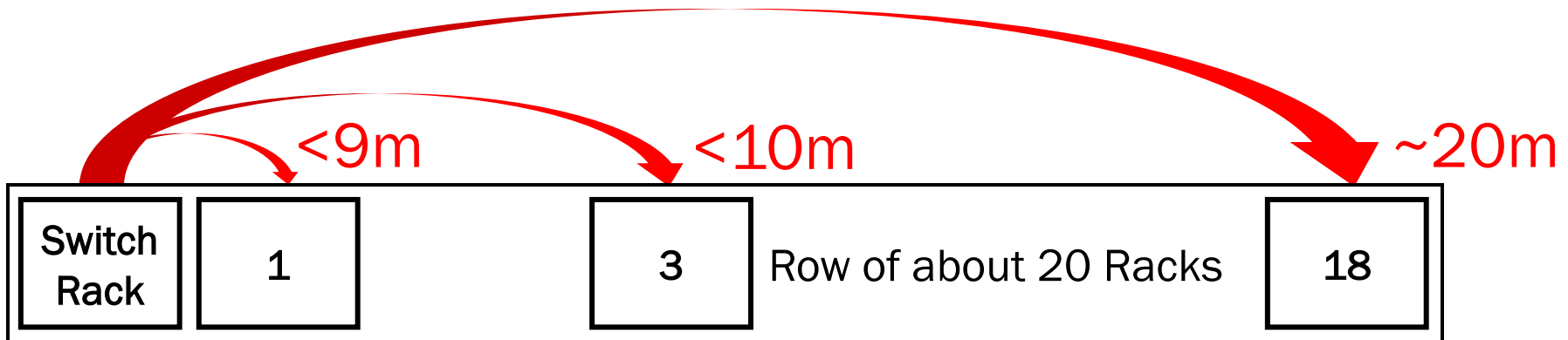


Link length from cable tray to any switch port in rack Switch = 4m

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

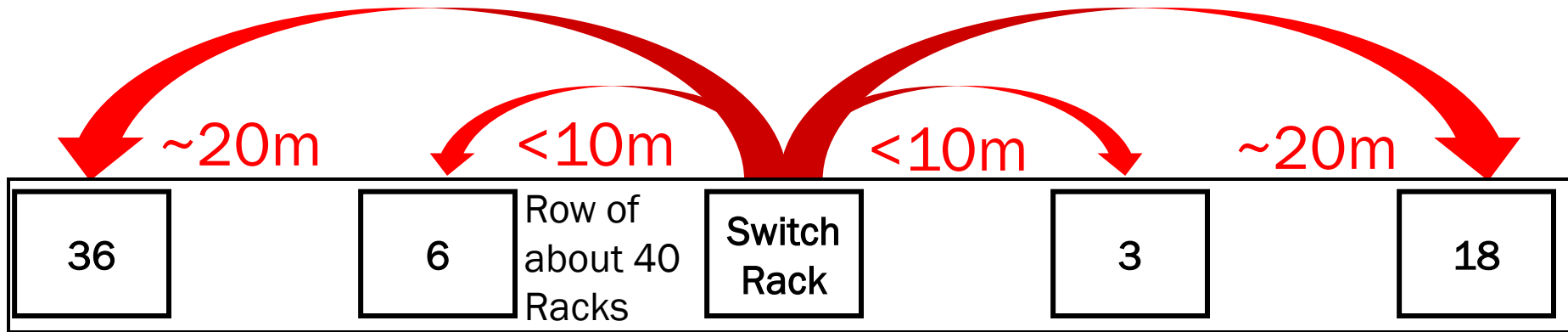
Reach 100s and ~1,000 servers with 20 meters



	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

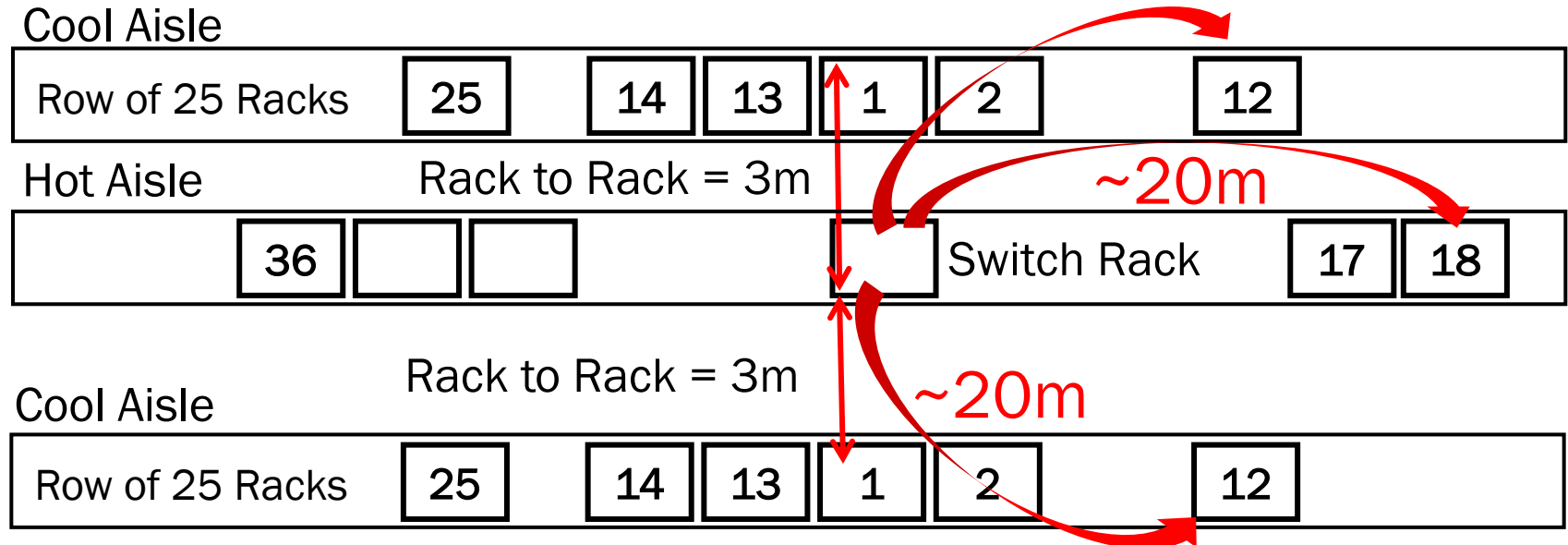
Reach 100s of servers with 10 meters, 1000s with 20m



	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

1,000 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

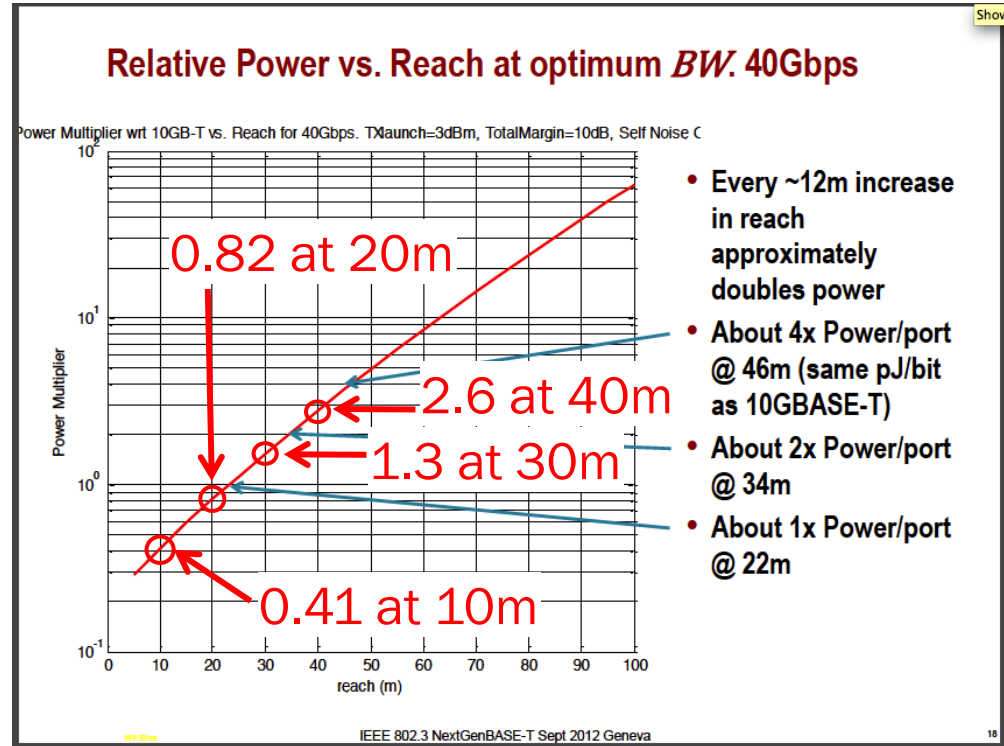


# Low Power or Low Density

- 10 meters = 0.4X 10GBASE-T Port
- 20 meters = 0.8X 10GBASE-T Port
- 30 meters = 1.3X 10GBASE-T Port

10GBASE-T Power	10m Power	20m Power	30m Power
2W	0.82W	1.64W	2.6W
3W	1.23W	2.46W	3.9W
4W	1.64W	3.28W	5.2W
5W	2.5W	4.1W	6.5W

2W/Port Power Line



bliss\_01\_0912 p18



# Bandwidth Length Product

- 10GBASE-T \* 100meters = 100Gbps\* meters
  - 3W now, still too high 10G at 100m
- 40GBASE-T \* 30 meters = 120 Gbps\* meters – Harder
  - 1.3X 10GBASE-T almost matches Bliss estimates 40G at 30m
- 40GBASE-T \* 20 meters = 80 Gbps\* meters – Easier
  - 0.8X 10GBASE-T correlates with Bliss estimates 40G at 20m
- 40GBASE-T \* 10 meters = 40 Gbps\* meters – Easiest
  - 0.4X 10GBASE-T correlates with Bliss estimates 40G at 10m
  - This could be a Short Reach limit to auto-negotiate down to



## 20 Meters or Less

- Low power (<2W) 40GBASE-T ports are desirable to yield a high density switch
  - 1U ToR switches only can support about 40 40GBASE-T ports, so very short distances of <10 meters are needed
  - Most modular switches will only support ~500 40GBASE-T ports in a 1 rack chassis until 2020
- 1,000 of high density, rack mounted servers can be supported with a 20 meter link to a modular switch
- I'll post my spreadsheet so you can modify assumptions as you would like
- 20 meters should be the reach of 40GBASE-T



**BROCADE** 

**THANK YOU**

