Design Challenges at Scale

HPC Users Group - Tucson
April 17, 2018

HPE leads the HPC market today

143 systems
29%of the Top500 List

TOP 500 The List.

World's most energy efficient supercomputer

THE GREEN 500

Largest Market Share
36 percent

HYPERION RESEARCH
The former IDC HPC analyst team

World's most secure industry standard server for HPC

HPE Apollo 6000 Gen 10

Fastest parallel processing performance

HPE SGI 8600

TSUBAME 3.0 at Tokyo Institute of technology

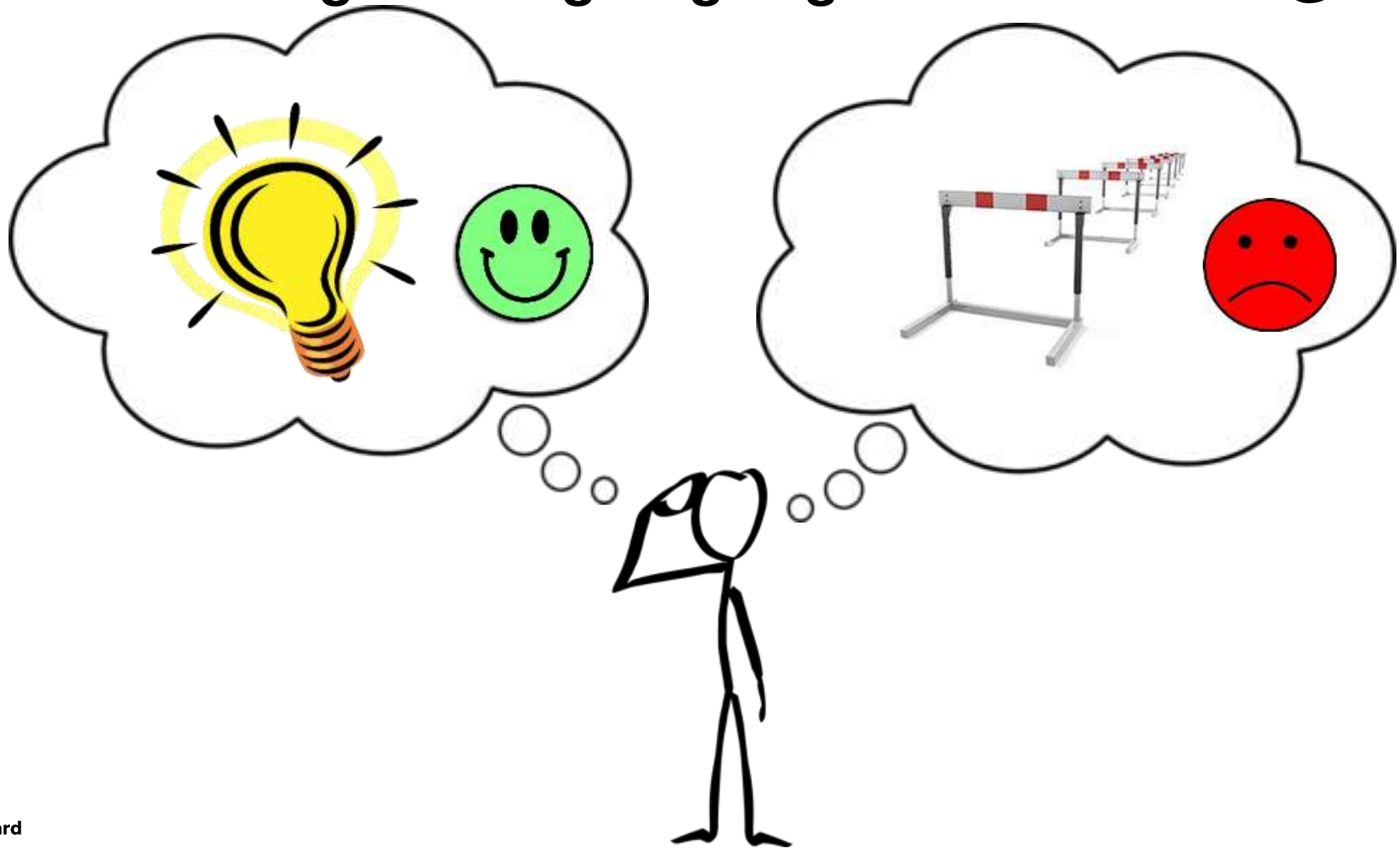
- Direct liquid cooling
- Petaflop scale HPC and AI
- #1 on the Green 500

First Memory-Driven Computing Prototype

the MACHINE

Hewlett Packard Enterprise

Several challenges designing High Performance @ Scale



Hewlett Packard Enterprise

Our Nemesis – the Active Optical Cable

(No offense intended to our AOC Partners)

Challenges

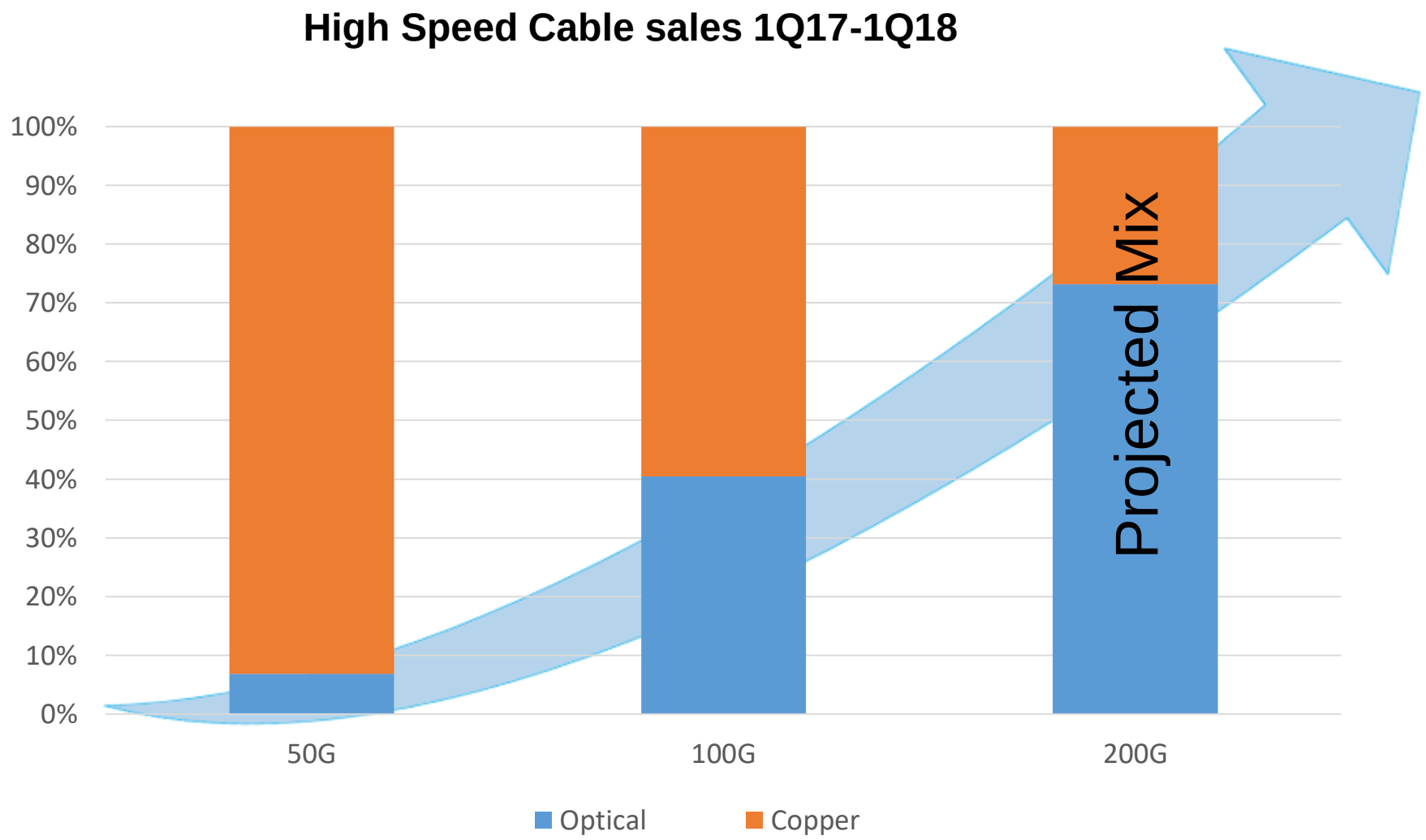
- Cost and Signaling
- Density and Egress
- Power and Thermals

Other opportunities

- “Wire Once” capabilities
- Affordable workload optimized topologies
- Reduce design dependencies on Radix

Hewlett Packard Enterprise

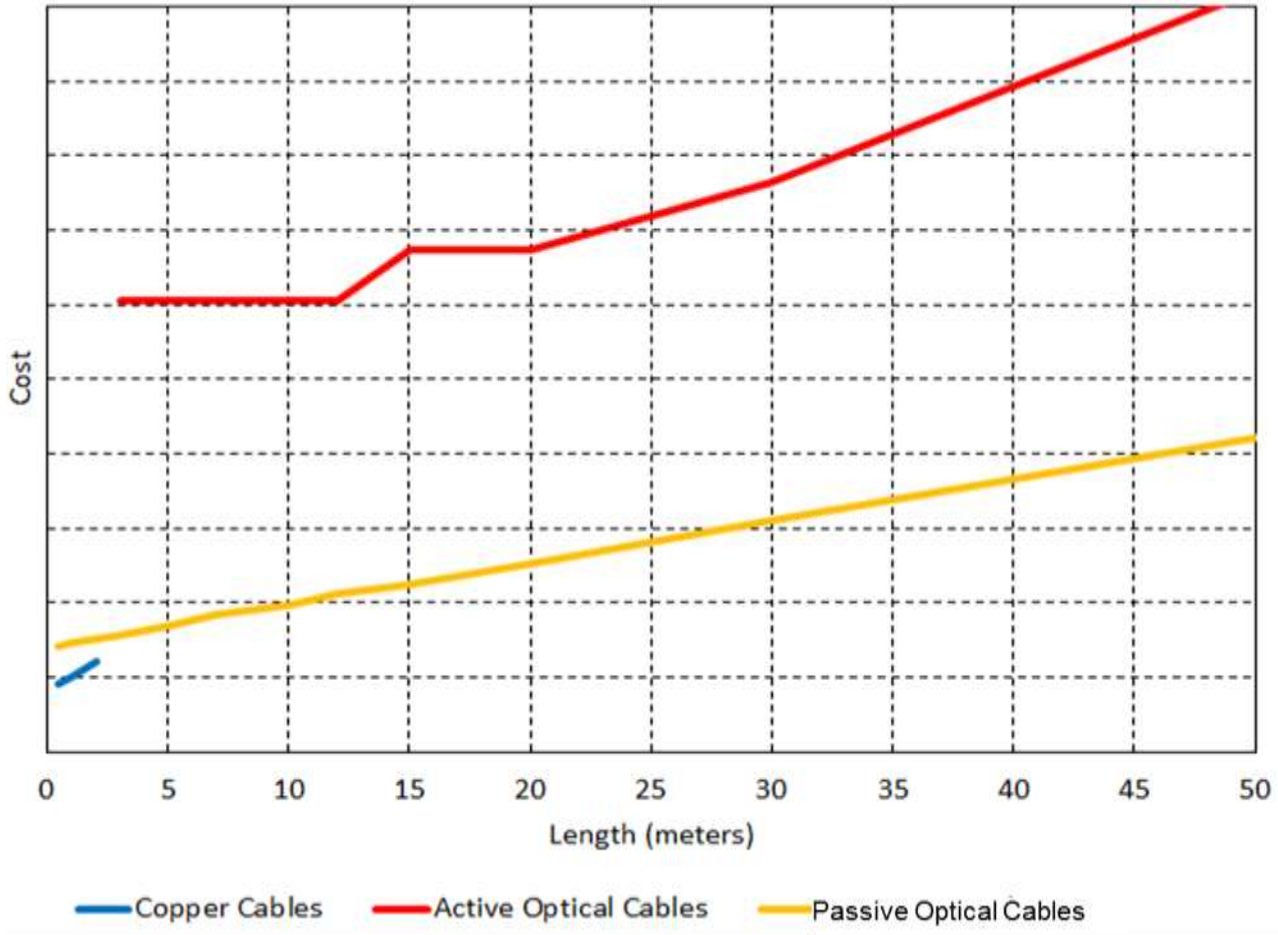
Optical Cable Sales continue to increase



Hewlett Packard Enterprise

Optical Costs driven by increased AOC complexity

Cable Cost Scalability



Estimated Average List Price

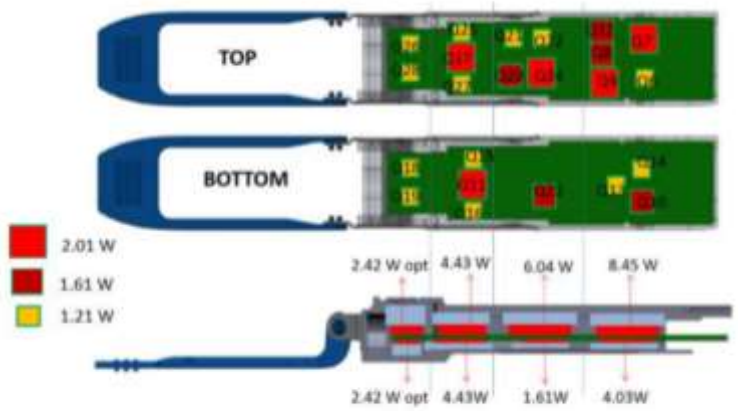
	50G	100G	200G	400G
Copper	$150	$250	$500	???
Optical	$620	$1000	$2150	OMG

The cost of the transport material

Copper = x

Glass = ~0.6x

The costs and complexity lie here!



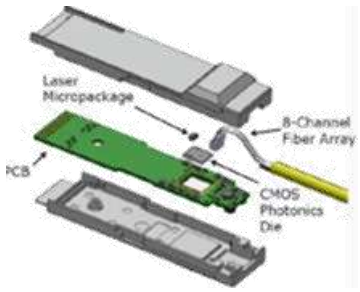
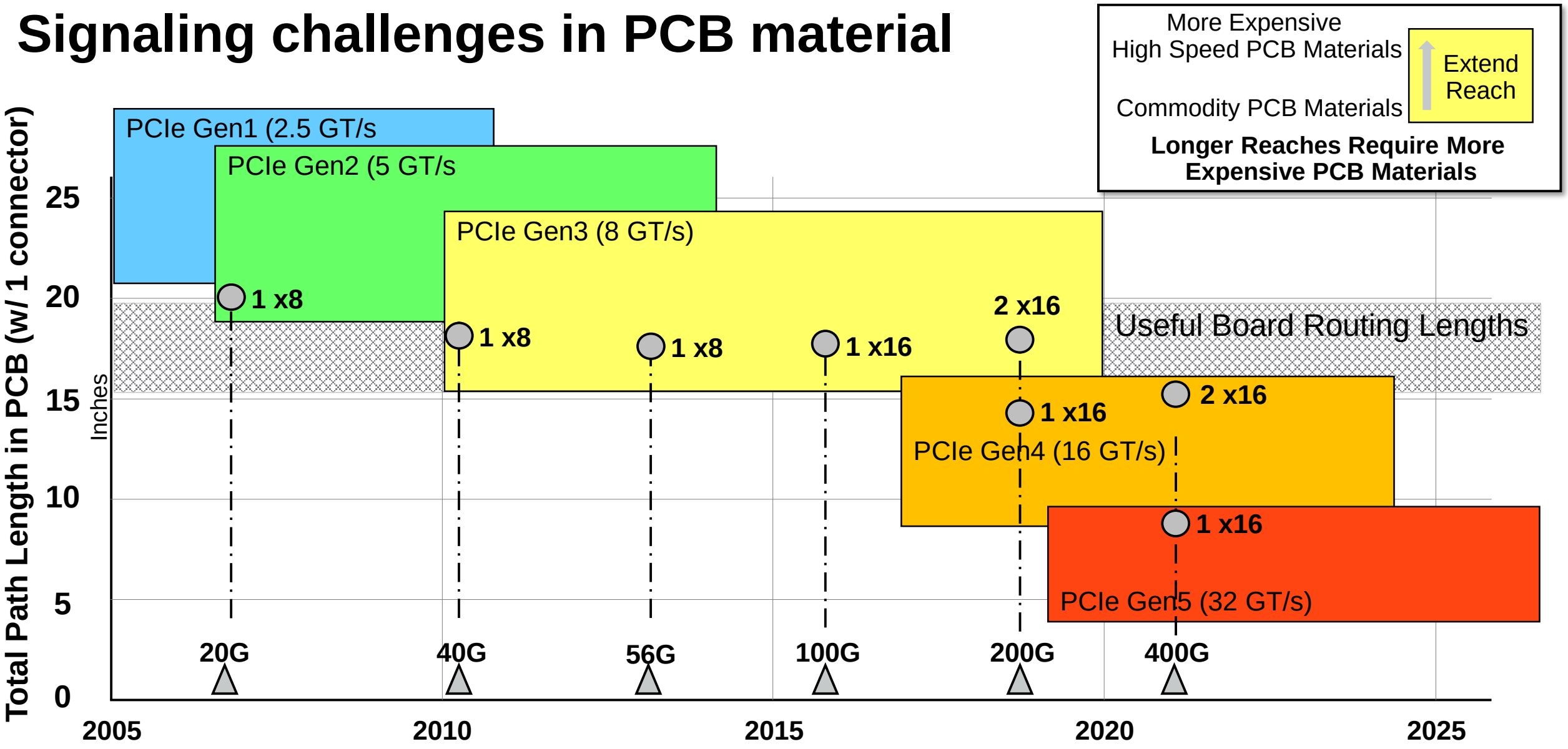


Figure 13: Heat Sources on Top and Bottom Side of PCB with Maximum Power Dissipation Capability per Region

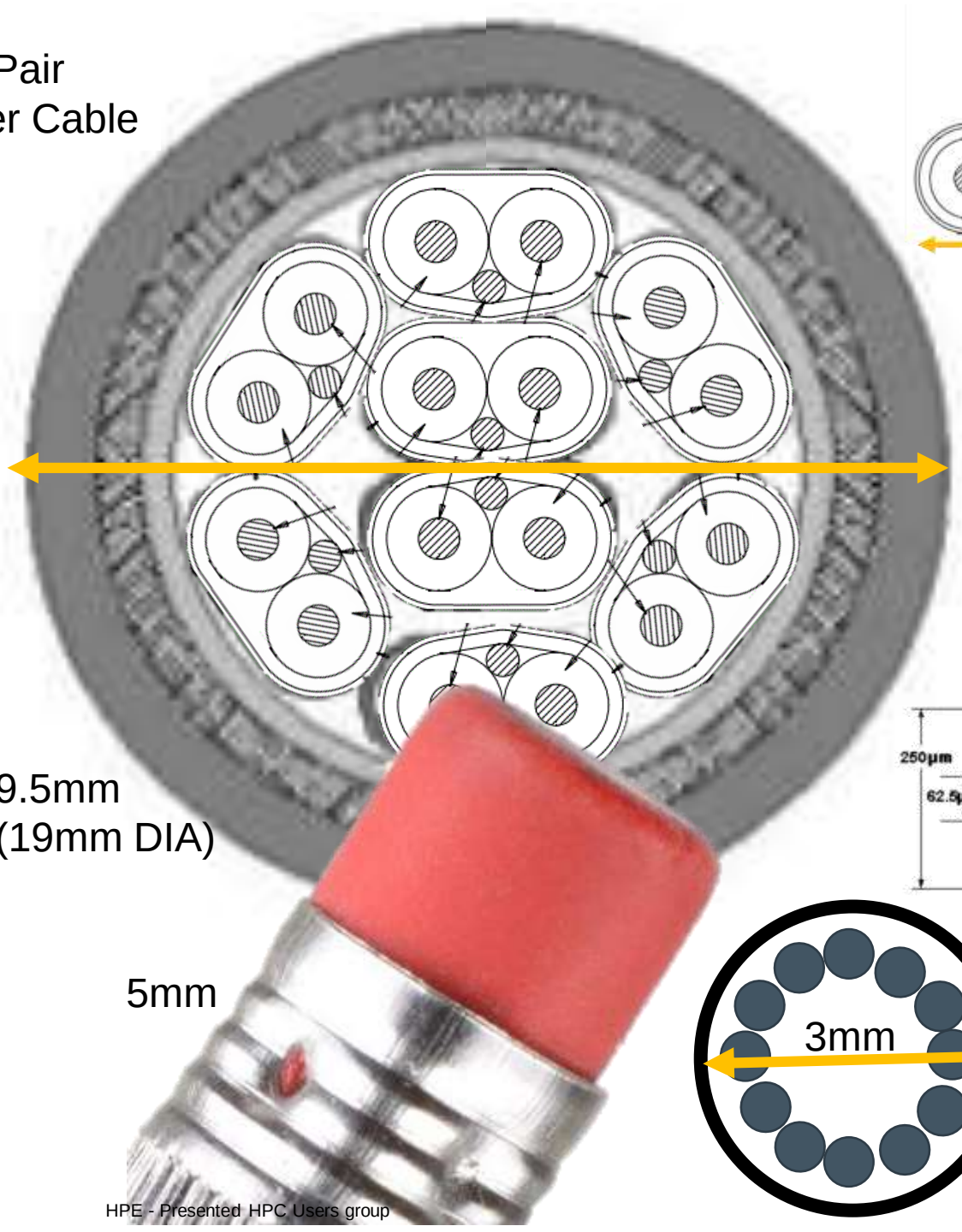
Hewlett Packard Enterprise

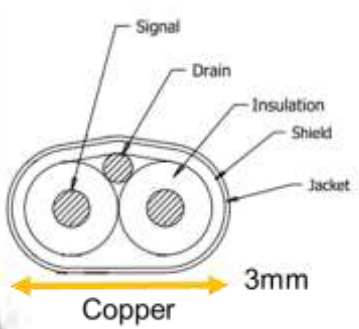
Signaling challenges in PCB material



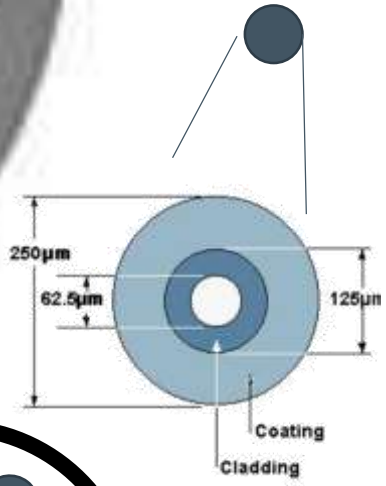
Hewlett Packard Enterprise

8 Diff Pair
Copper Cable





Optic Fiber
0.25mm

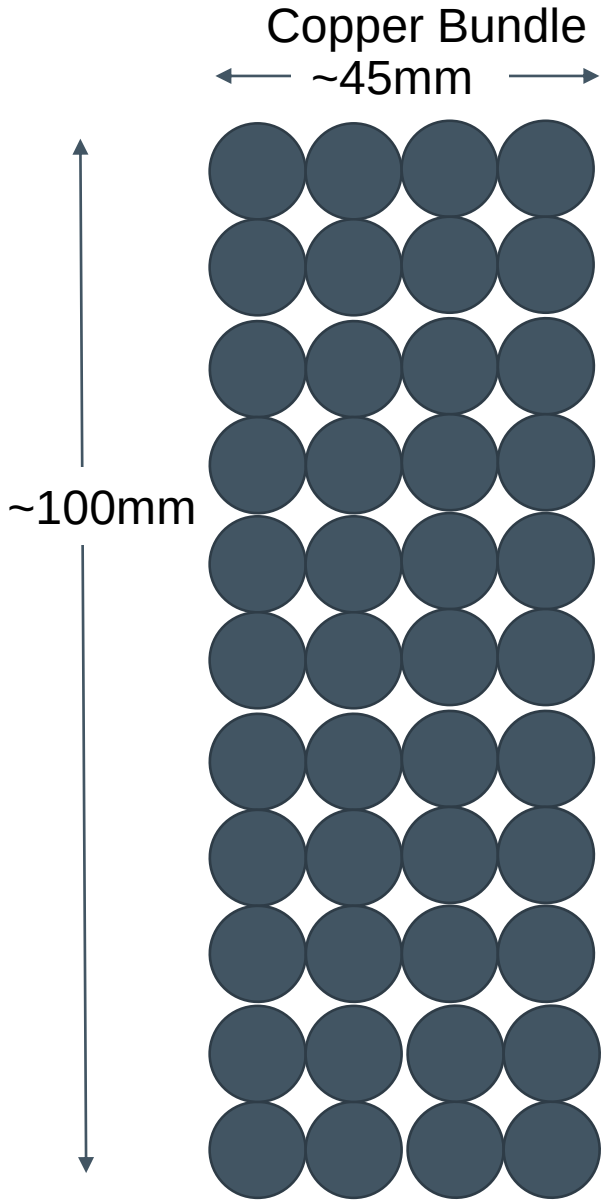


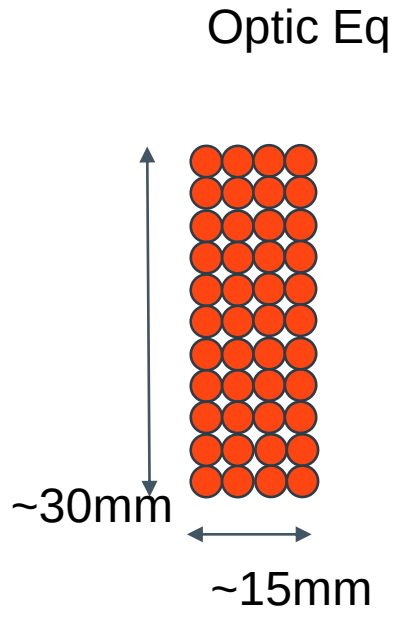
9.5mm
(19mm DIA)

5mm

12 Fiber
Optic Cable

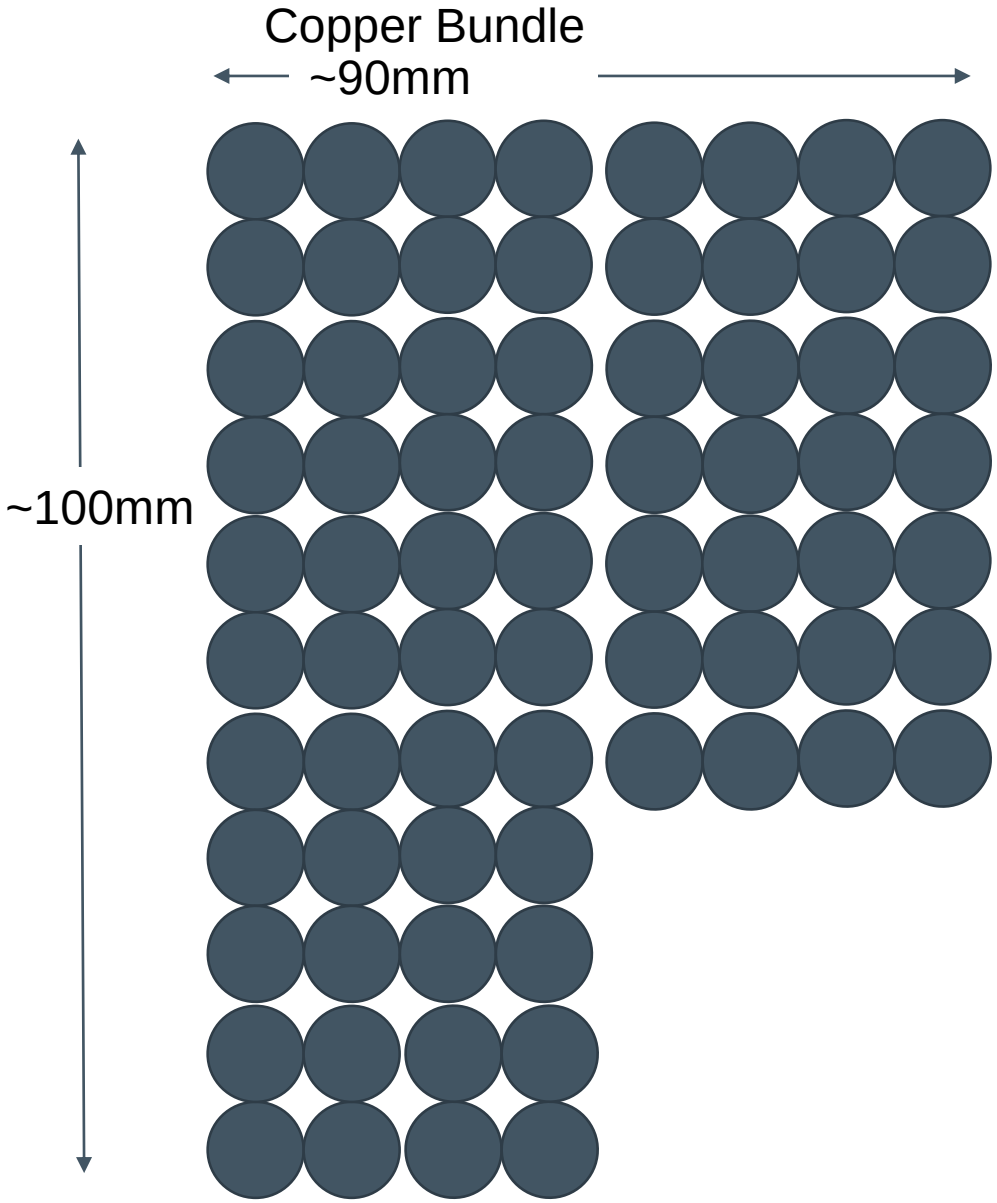
72 Node Rack 1:1 Fat Tree – Single Side of Rack

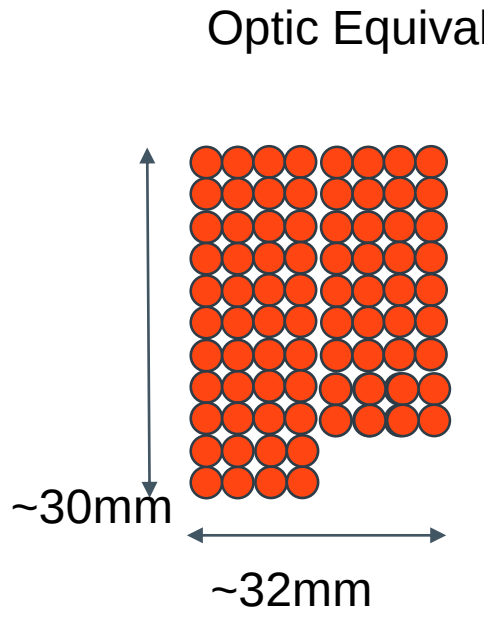




Hewlett Packard Enterprise

128 Node Rack 1:1 Fat Tree – Single Side of Rack





Hewlett Packard Enterprise

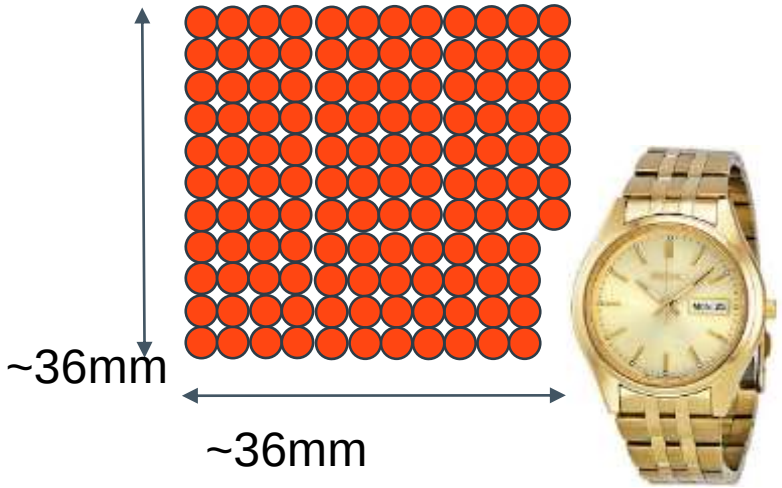
256 Node Rack 1:1 Fat Tree – Single Side of Rack

Copper Bundle

~120mm

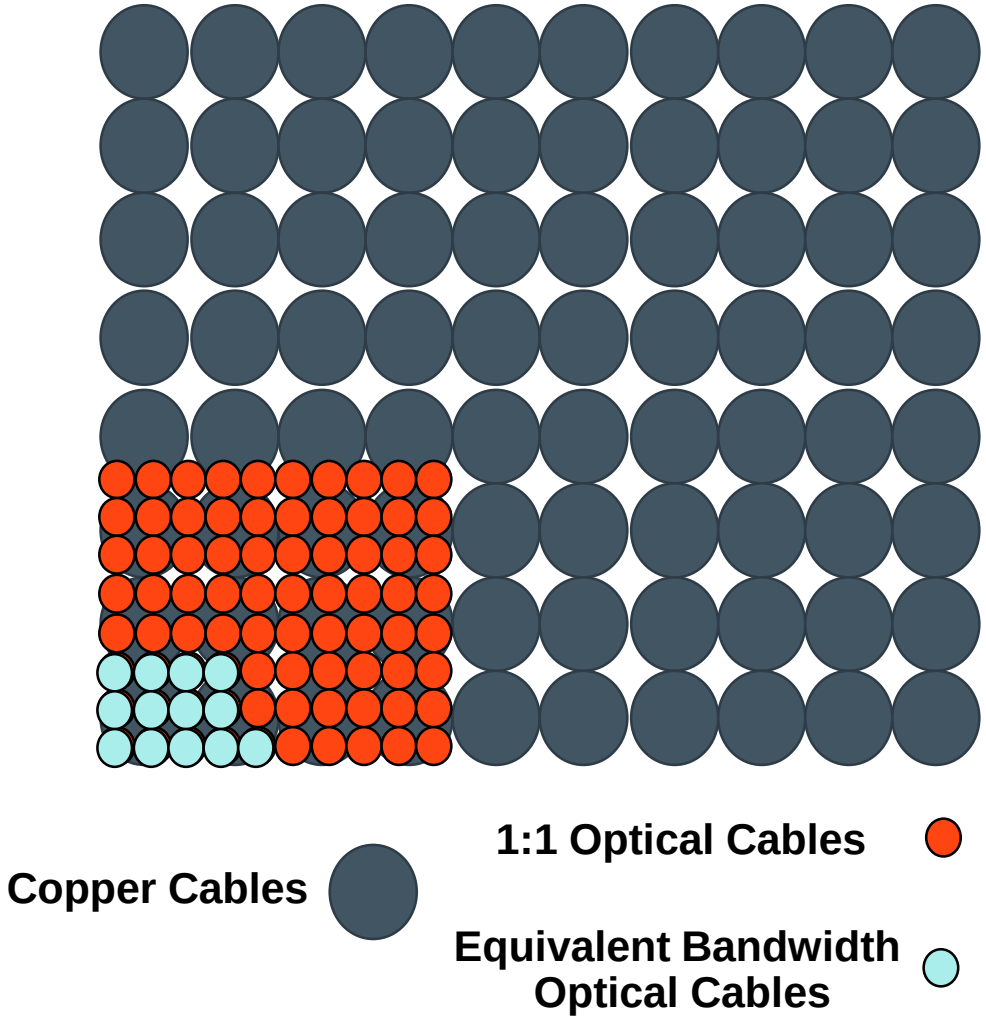
~120mm

Optic Equivalent



Hewlett Packard Enterprise

Copper becomes a challenge at scale

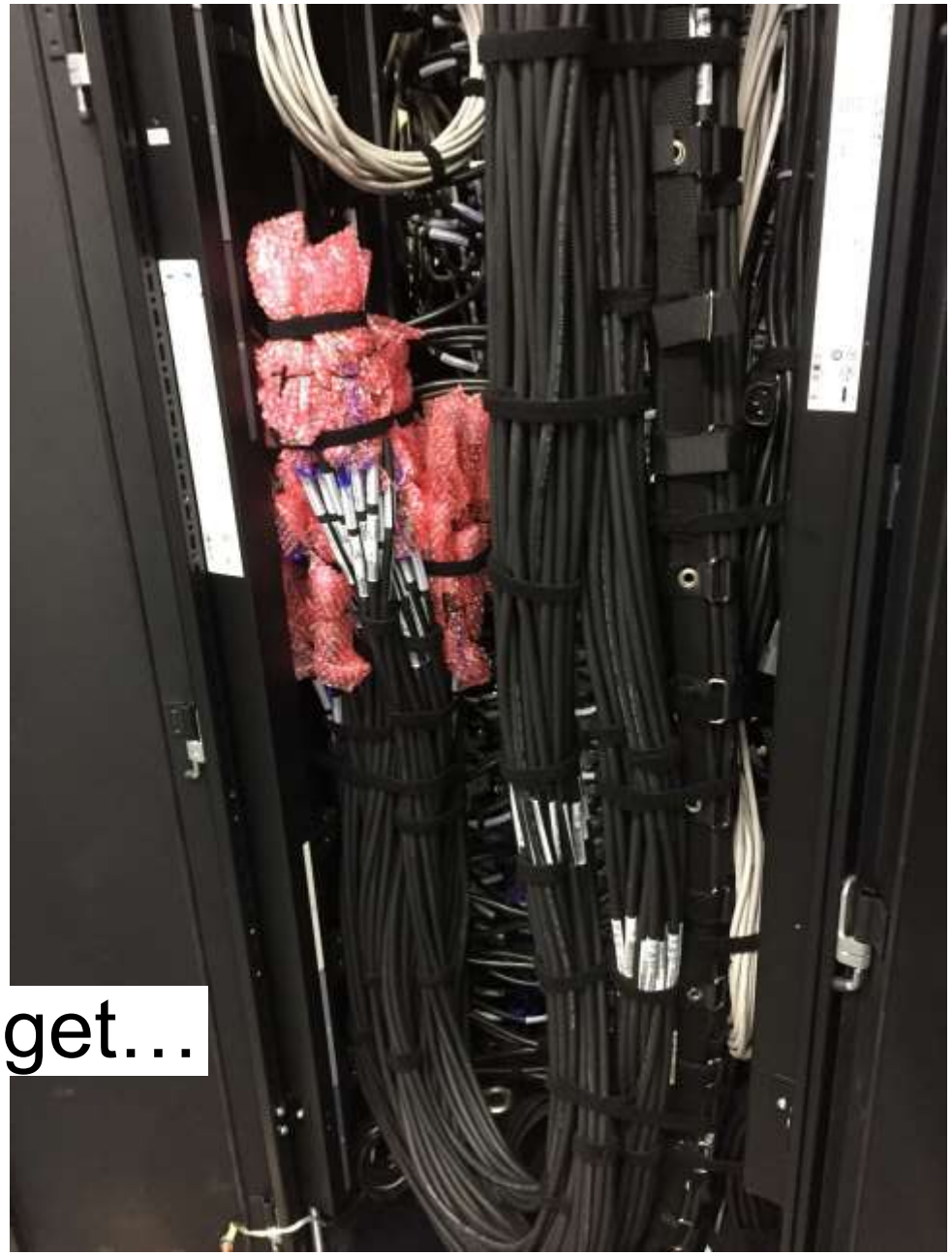


Hewlett Packard Enterprise

Egress

When you put several of these in a rack…

Apollo 6000 OPA Switch



This is what you get…

Hewlett Packard Enterprise

Power and Thermals

– Optical Cable adoption is driving increased power in the switch

Approximate Cable Power

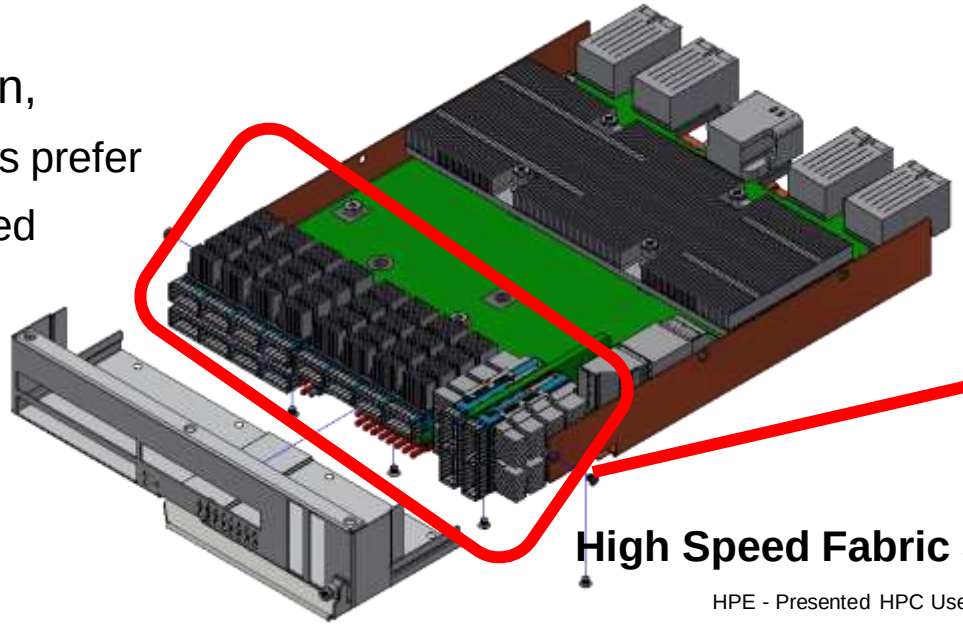
Copper	50G	100G	200G	400G
0W	1.5W	2.2W	5W	??W

– This drives a commensurate thermal issue in the switch

– In System Design,

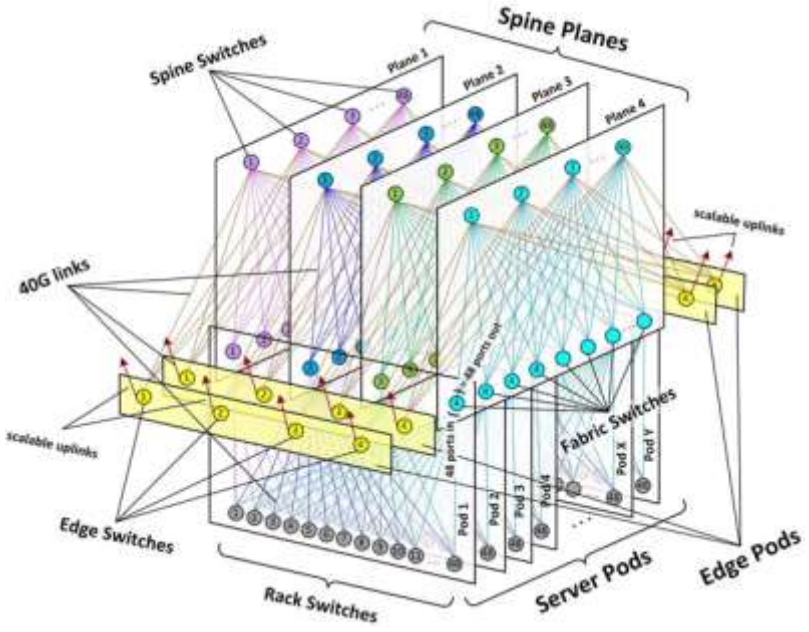
Cooling techniques prefer planar and localized

This is NEITHER

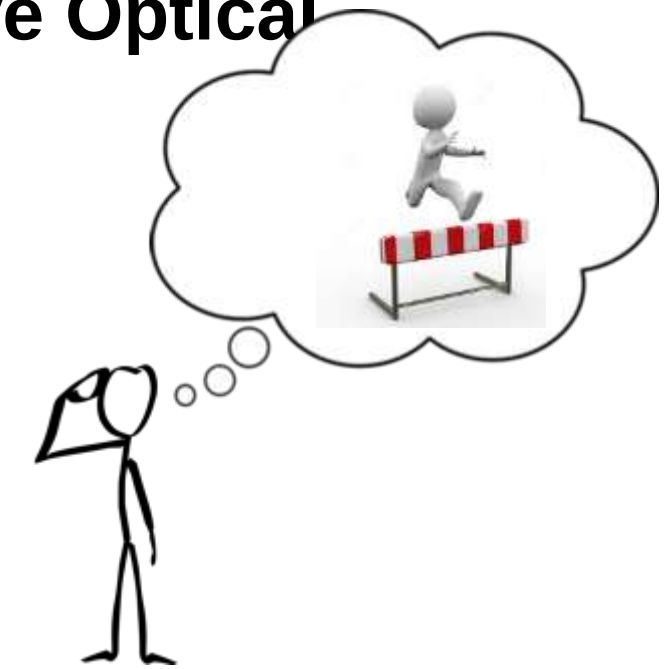


High Speed Fabric Switch

Other Degrees of Freedom if we moved to Passive Optical

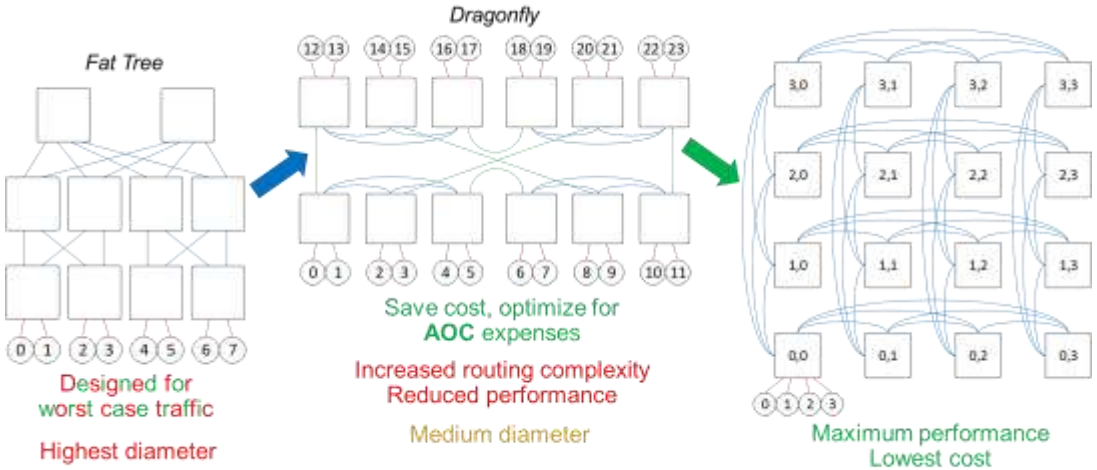


Wire Once Capability

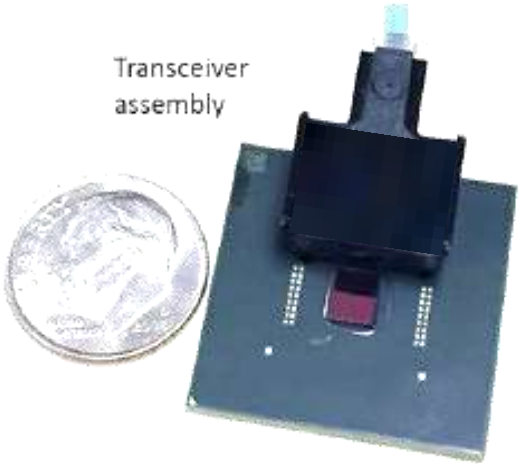


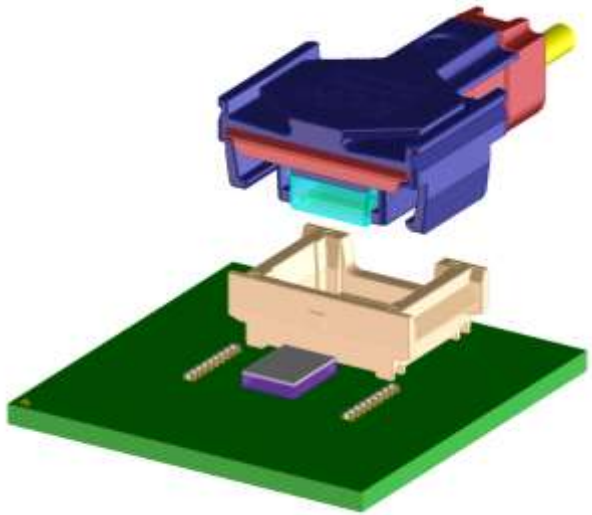
Design relief to drive to more physical constraints

Optimize fabric topologies for workloads not cable costs



The time for affordable, reliable mid-board optics is NOW!!



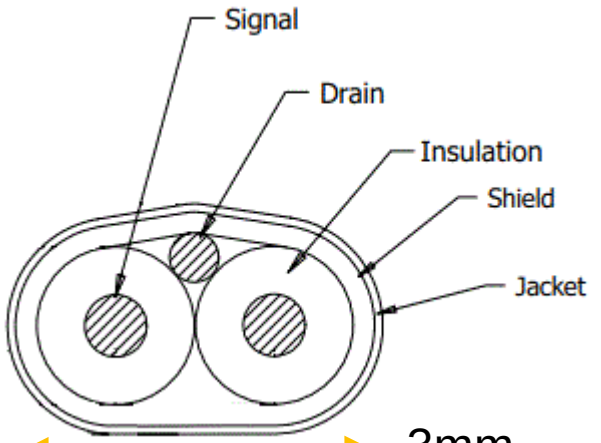


Hewlett Packard Enterprise

Thank You!

Hewlett Packard Enterprise

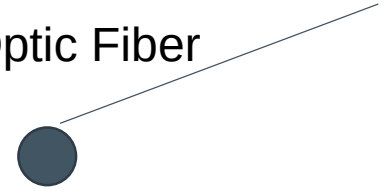
9.5mm
(Penny is 19mm diameter)



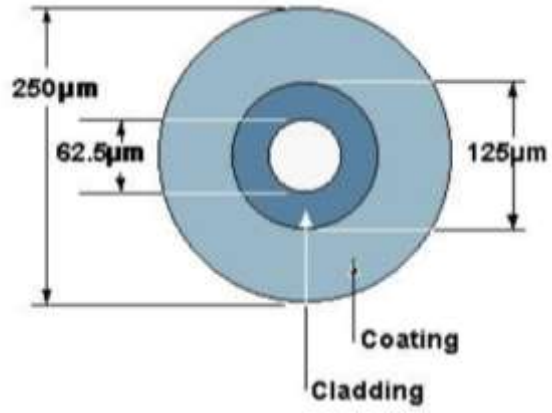
3mm

Copper

Optic Fiber



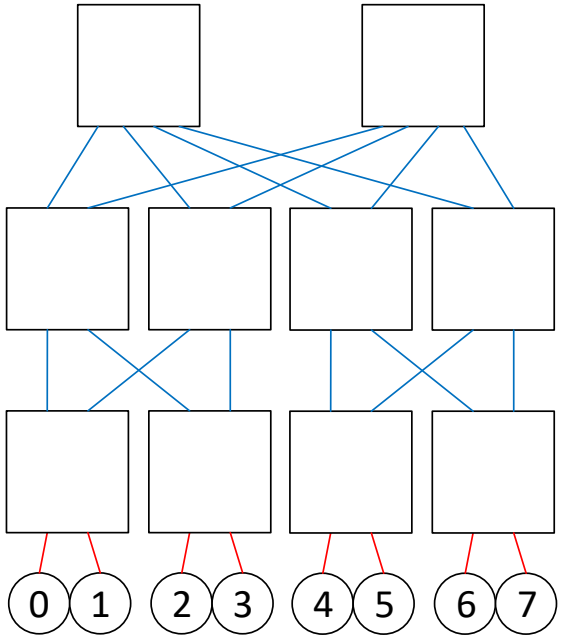
0.25mm



TYPICAL MULTIMODE CROSS-SECTION

We can stop optimizing our networks for cable costs

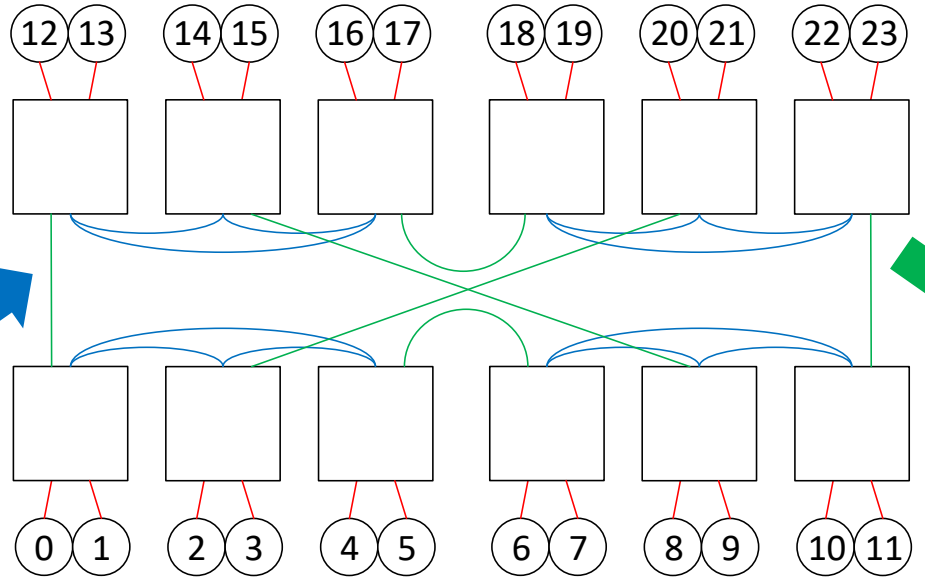
Fat Tree



Designed for
worst case traffic

Highest diameter

Dragonfly

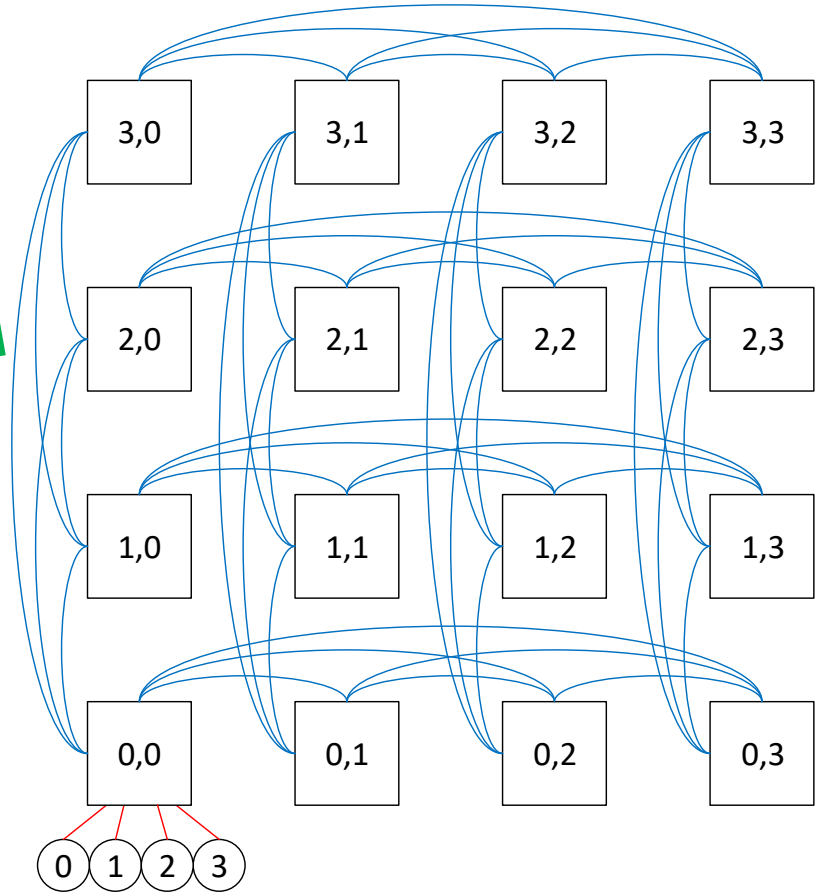


Save cost, optimize for
AOC expenses

Increased routing complexity
Reduced performance

Medium diameter

HyperX



Maximum performance
Lowest cost
Lowest diameter

Hewlett Packard Enterprise