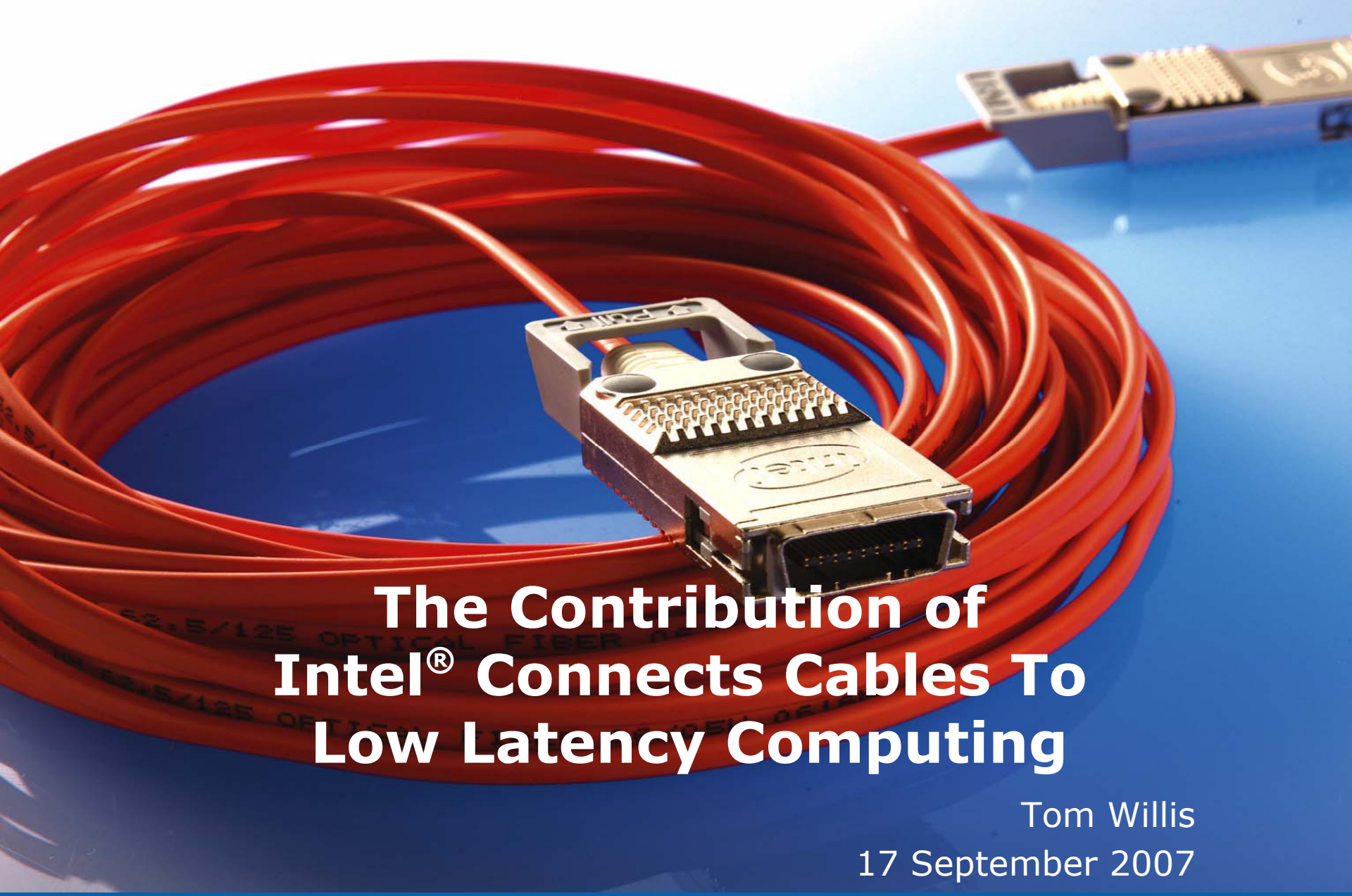


The Contribution of Intel® Connects Cables To Low Latency Computing

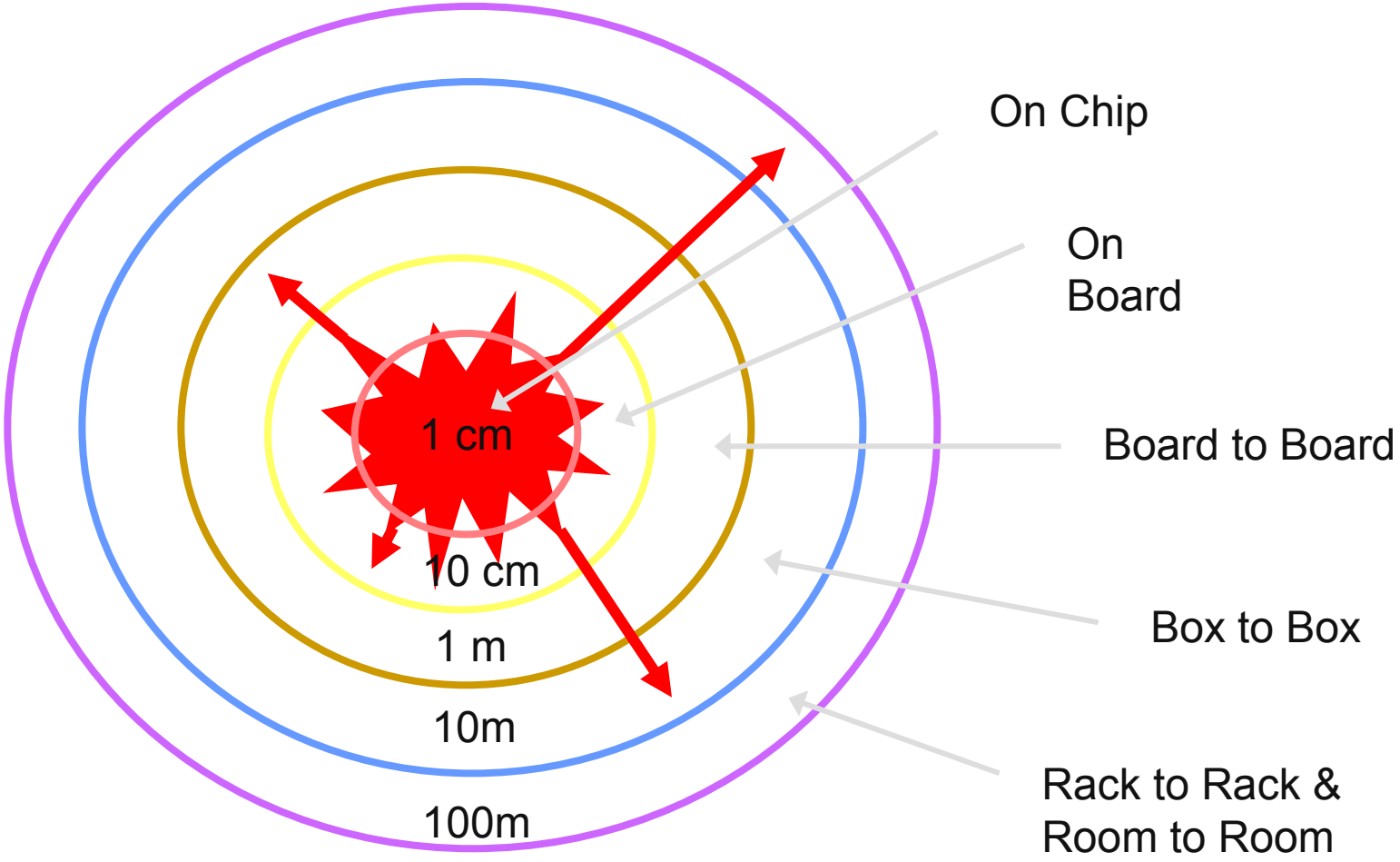
Tom Willis

17 September 2007

Agenda

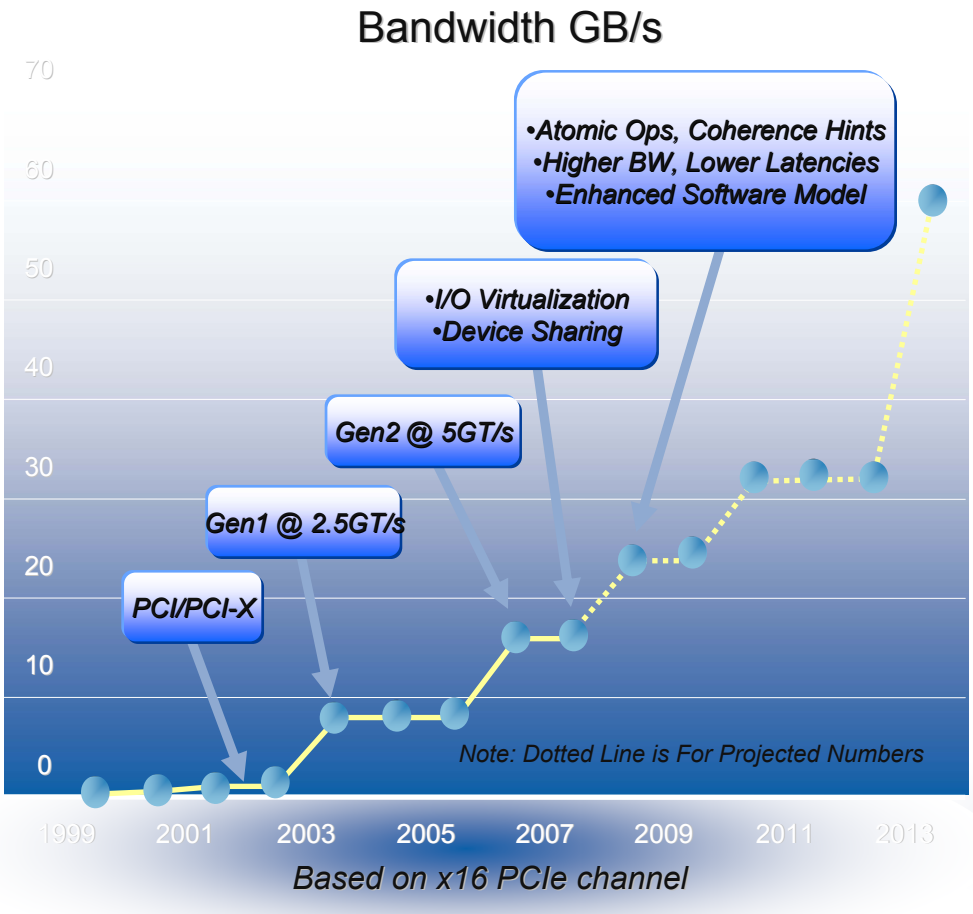
- The Data Rate Explosion
- Cable effects on latency
- Scaling low latency systems

Higher Data Rates Are Exploding Out From the Microprocessor

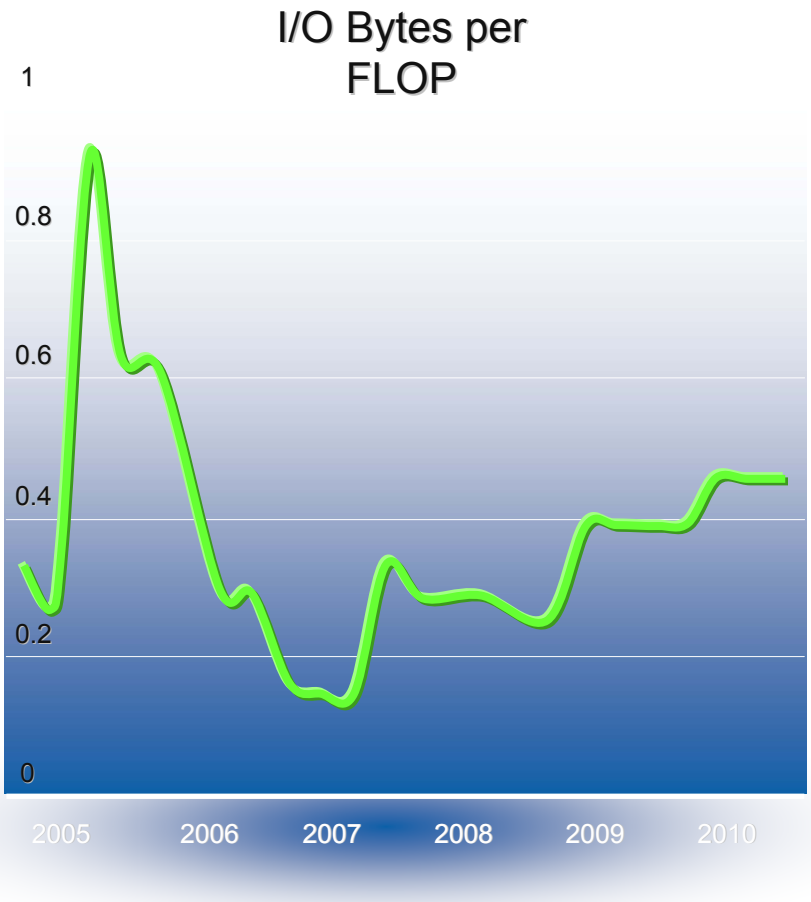


PCI Express to Meet I/O Demand

Performance, Bandwidth and Functionality



Based on x16 PCIe channel



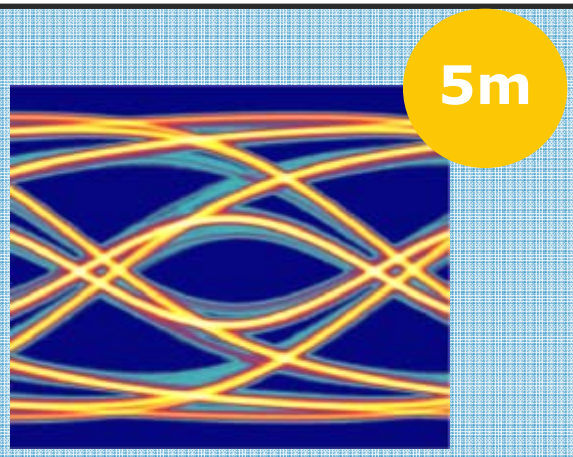
Tracking Moore's Law

Source: Intel

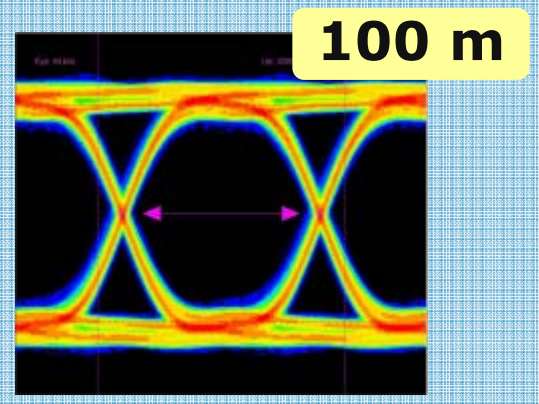
Copper Interconnects Are Struggling

- Copper signal quality decreases as data rates and distance increase
  - Required power
  - Electro Magnetic Interference
  - Weight
  - Latency
- Optical Signal Quality is orders of magnitude better
  - Longer wavelength
  - No charge effects

5 Meter 24 AWG Copper Cable



100 Meter Intel® Connects Cable



One Example – 10GbaseT Versus Intel Connects Cables

- 10 watt disadvantage per link
  - >10GBaseT *target* power of 6 watts per side
  - >Intel Connects Cable power today = ~1 watt per side
- ~2000 times more conversion latency
  - >10GBaseT *target* conversion latency of 2 microseconds per link
  - >Intel Connects Cable is ~550 picoseconds today
- 1000 times more bit errors
  - >10GBase T *target* Bit Error Rate is 10^{-12}
  - >Intel Connects Cable is 10^{-15} BER today
- ~2X greater volume and weight

Agenda

- Context
- Cable effects on Latency
  - Simple cable latency
  - Effective latency
- Scaling low latency systems

Simple Cable Latency for Intel Connects Cables

- Optical/Electrical Conversions = 0.275 nanoseconds
- Speed of light through the fiber = 4.99 nanoseconds per meter
- Latency of a 10 meter Intel Connects Cable
  - First O/E conversion: 0.275
  - Speed of of light 10m*4.99 ns: 49.9
  - Second O/E conversion: 0.275
  - Total: 50.45 ns
- Latency of a 100 meter Intel Connects Cable
  - First O/E conversion: 0.275
  - Speed of of light 100m*4.99 ns: 499.0
  - Second O/E conversion: 0.275
  - Total: 499.45 ns

Effective Latency

- Function of
  - Simple Cable Latency
  - Bit Error Rate
  - The time required to find and fix those bit errors
    - Many things affect this
      - Other physical delays (e.g. passing through switches)
      - Where the error is detected
      - Is it the Bit Error random, or is there a bad link ?
      - Whether the data to be resent is in a buffer or has to be re-accessed from slower media.
      - The system and application tolerance for bit errors.

Bit Errors Can Be Significant @ 20Gbps

Bit Error Rate at 20 Gbps per link

	10^{-12} BER @20Gbps	10^{-15} BER @20Gbps
Errors Per Day for a single link	1728	1.7
Errors Per Day For 1000 links	1,728,000	1,728

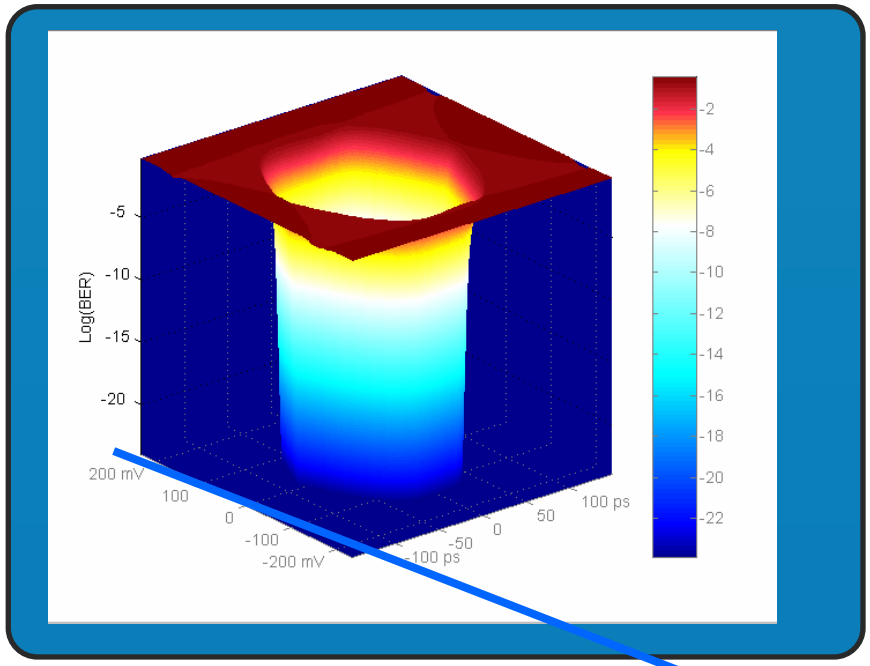
10^{-12} interconnects BER/day for 1000 links

Intel® Connects Cable BER/day for 1000 links

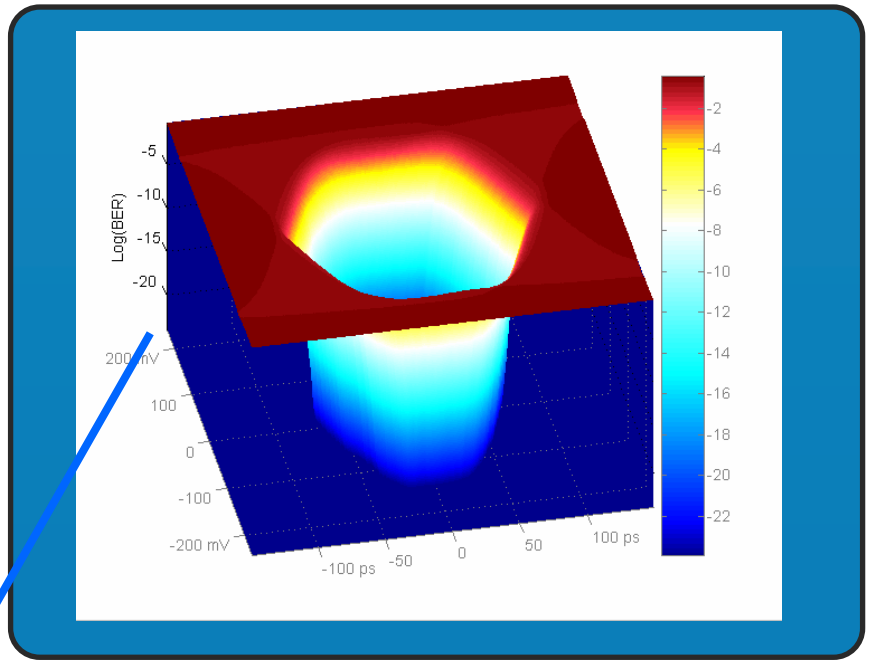
1000 times less BER than interconnects at 10^{-12}

Intel® Connects Cables
Actual Bit Error Rates May Even Be Lower*

10 Meter Intel® Connects Cable**



100 Meter Intel® Connects Cable**



10^{-25}

Extremely low BER for high HPC compute fabric stability

*Note: Specified BER for Intel® Connects Cables is 10^{-15}

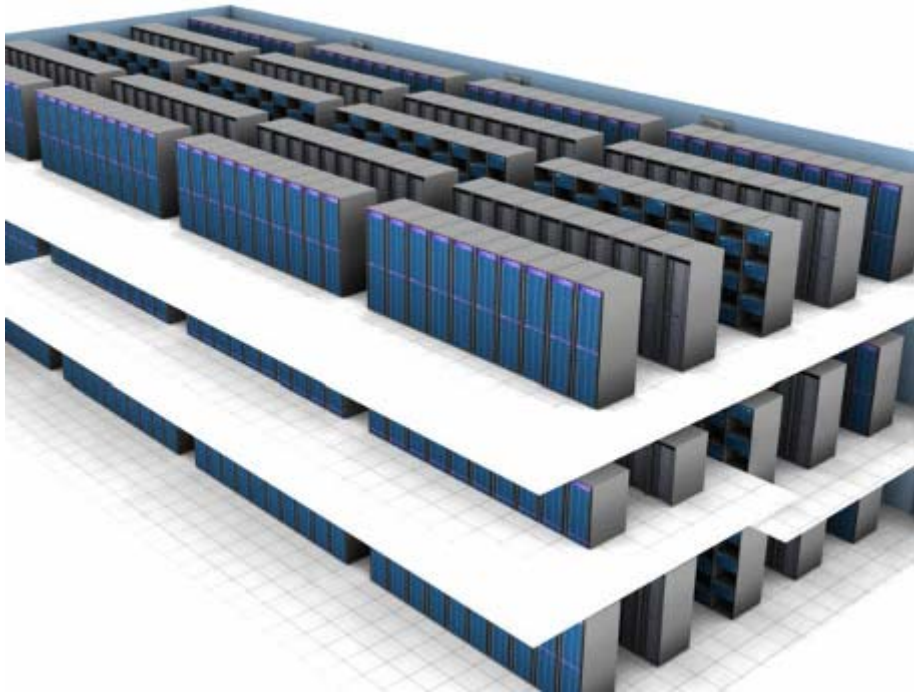
**Source: Tektronix Lab Evaluation

Agenda

- Context
- Cable effects on latency
- Scaling low latency systems

Intel® Connects Cables

Longer: 100m for Cluster Scaling



- **Build larger clusters with 20 Gbps DDR**
- **Design clusters based on needs not on cable lengths**

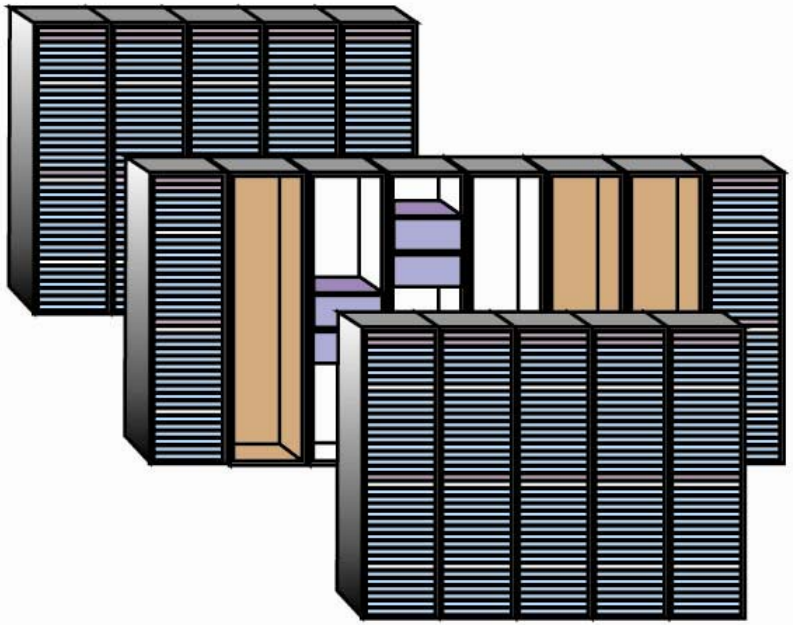
10X longer than 24 AWG copper cables**

*Fat Tree Architecture

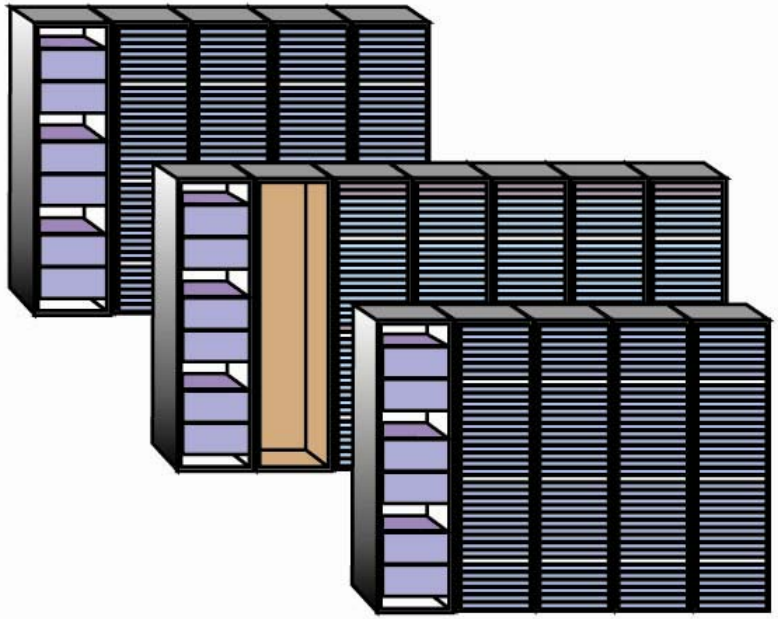
**for double data rates

Modified Traditional Design

Traditional design with short copper cables
Switches and servers interspersed, with space for bulky copper cables

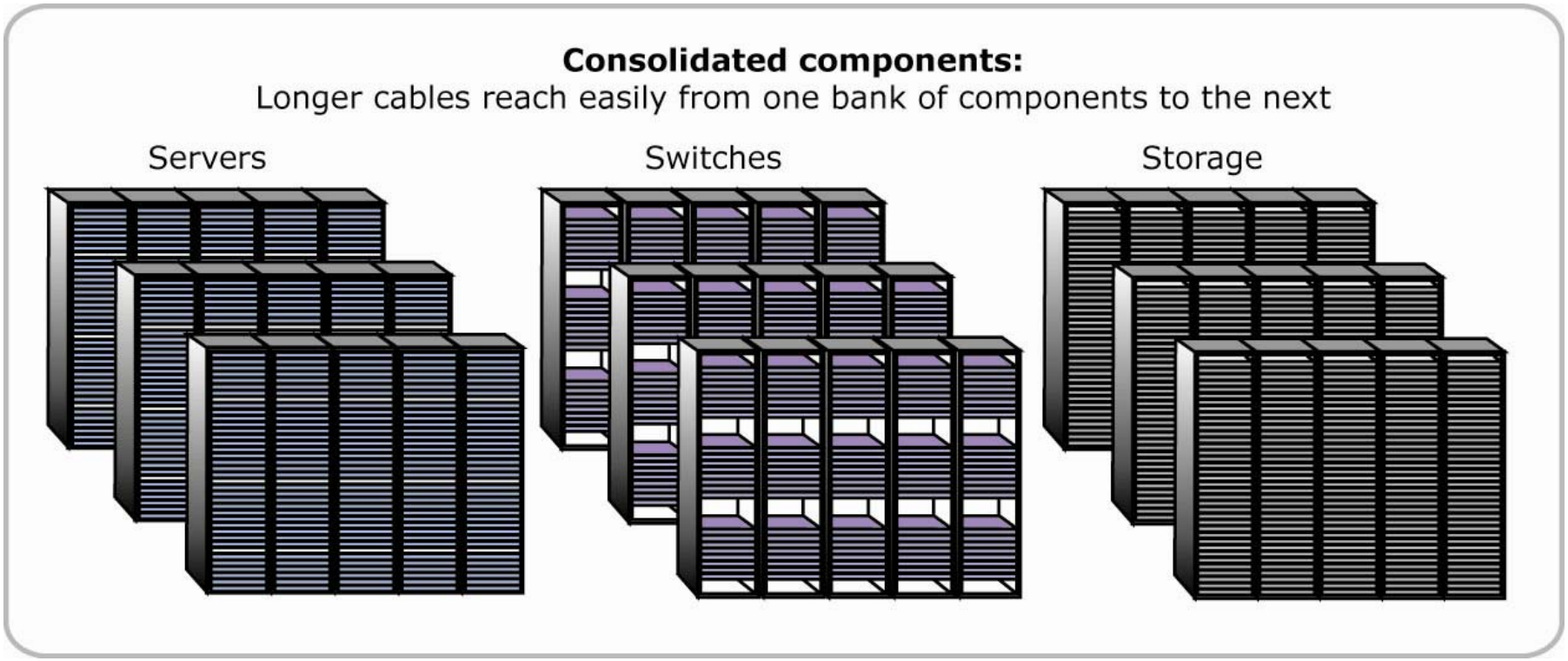


Modified traditional design with long optical cables
Switches grouped at end of row; eliminate "wasted" cable cabinets



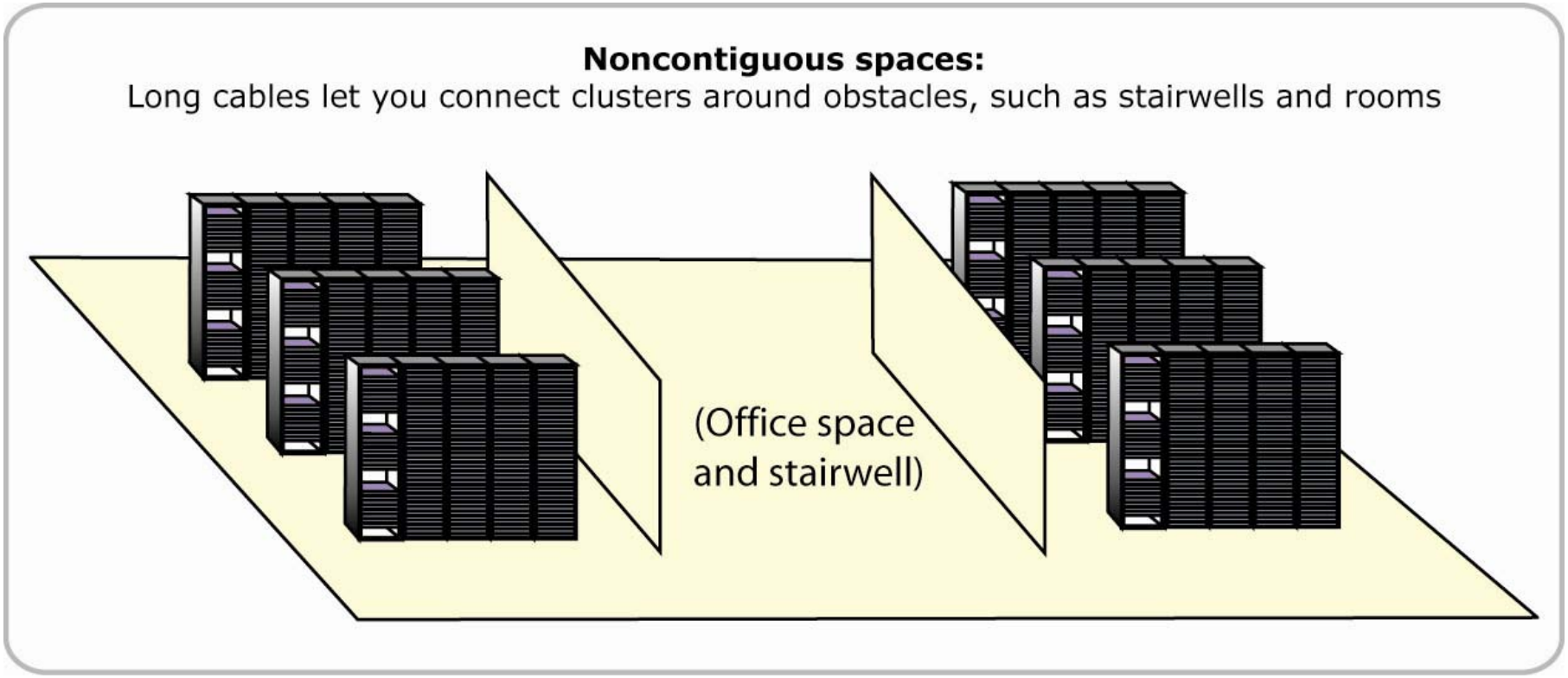
Longer, thinner cables let you move switches to the end of the row, eliminate excess cable cabinets, and increase compute density

Consolidated Components



Longer cables let you group components for easier installation, maintenance and upgrades, and more specialized cooling

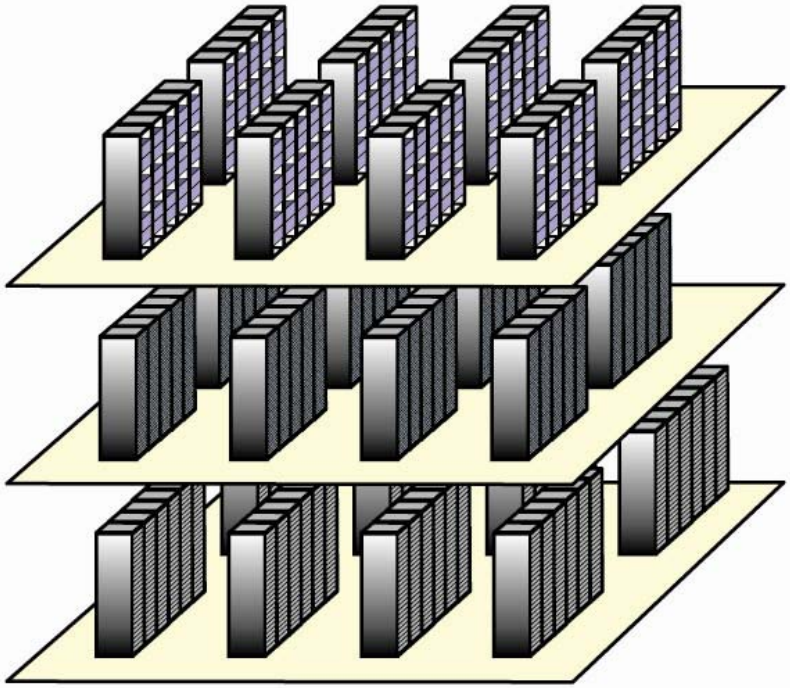
Non continguous spaces



Long cables let you design for the spaces you have, around pillars and stairwells, and across noncontiguous spaces, such as offices and meeting rooms.

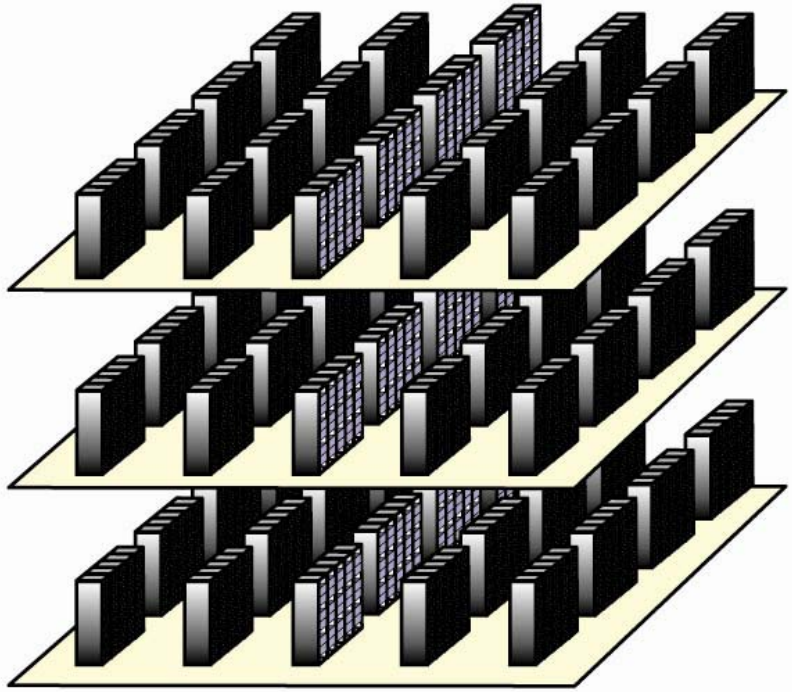
Vertical models

Vertical model: cake layer



Group components by floor to make cooling and maintenance easier

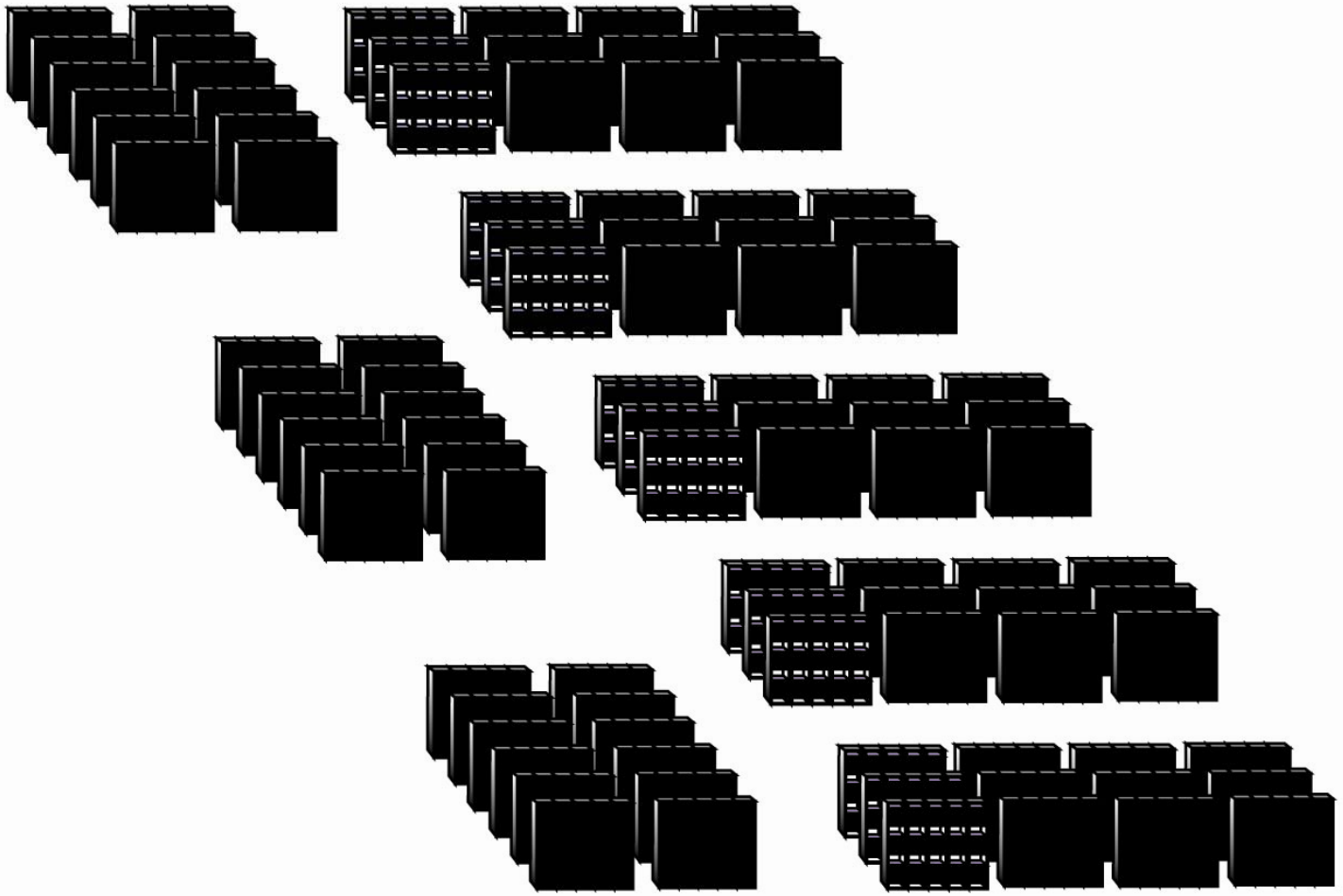
Vertical model: cylindrical



Build a cluster on each floor so that adding new clusters can be done floor-by-floor

Very Large HPC Facility

Very large HPC cluster:
Long cables let you build scalable units (smaller clusters) and connect them over large distances



Intel® Connects Cables

Lighter: The 10 Meter Challenge

Intel® Connects Cable **24 AWG Copper Cable**

.18 Kg

1.135 Kg

- Less weight to bend pins on servers
- No need to reinforce floors and cable ladders

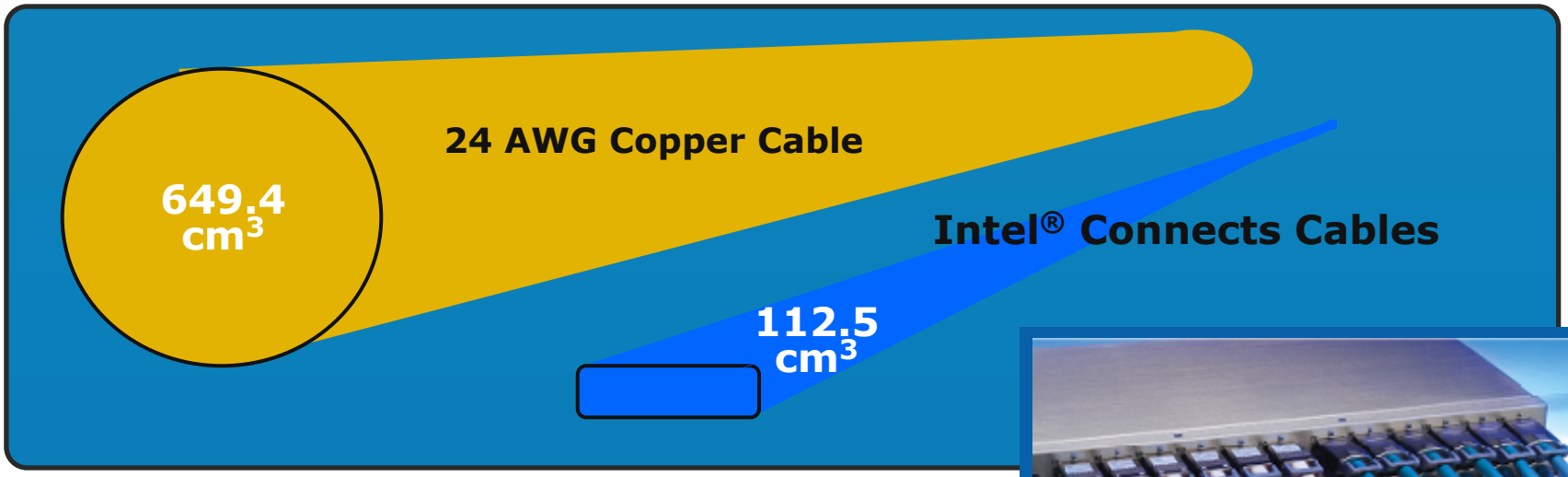
84 % lighter than 24 AWG copper cables*

*Source: Intel internal testing

Intel® Connects Cables

Thinner: Better Airflow in Racks, Floors

Volume Comparison of 10 m Cables*

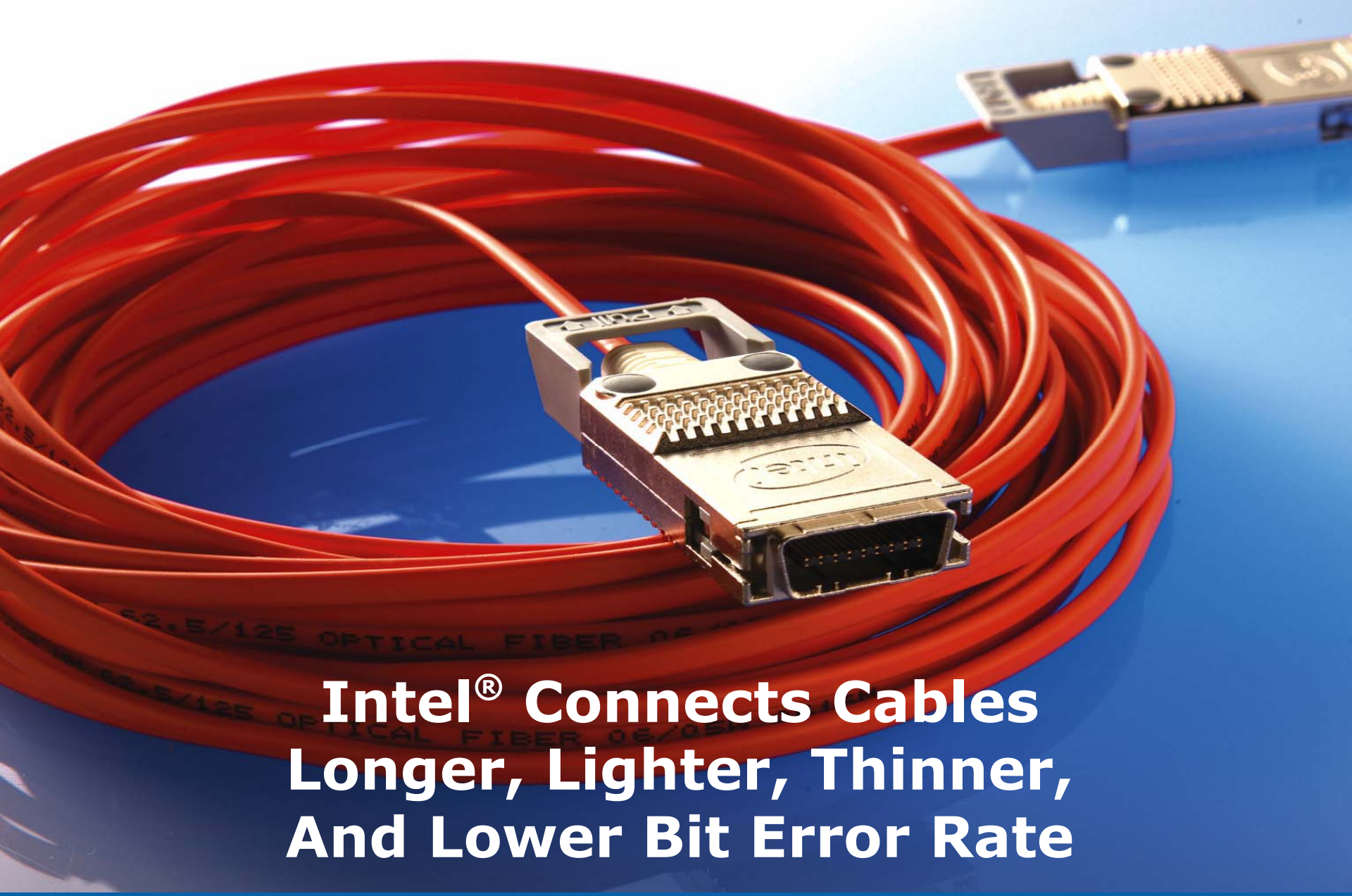


Comparison of 12 Intel® Connects Cables to 12 24-AWG copper cables

83% smaller than 24 AWG copper cables**

*Does not include connectors

**Source: Intel internal testing



Intel® Connects Cables Longer, Lighter, Thinner, And Lower Bit Error Rate