10GBASE-T: Broad Market Potential in Data Centers

Mike Bennett

11/7/2003

Contributors/Supporters

- Brent Draney – National Energy Research Supercomputing Center
- Roberto Morelli – Energy Sciences Network
- Greg Chartrand – Pacific Northwest National Laboratory
- Mike Bennett – Lawrence Berkeley National Laboratory

Discussion

- Where does 10GBASE-T fit in the market?
- What are the drivers?
- What are the alternatives?
- Data Centers
- Questions
- Conclusion

Where does 10GBASE-T fit in the market?

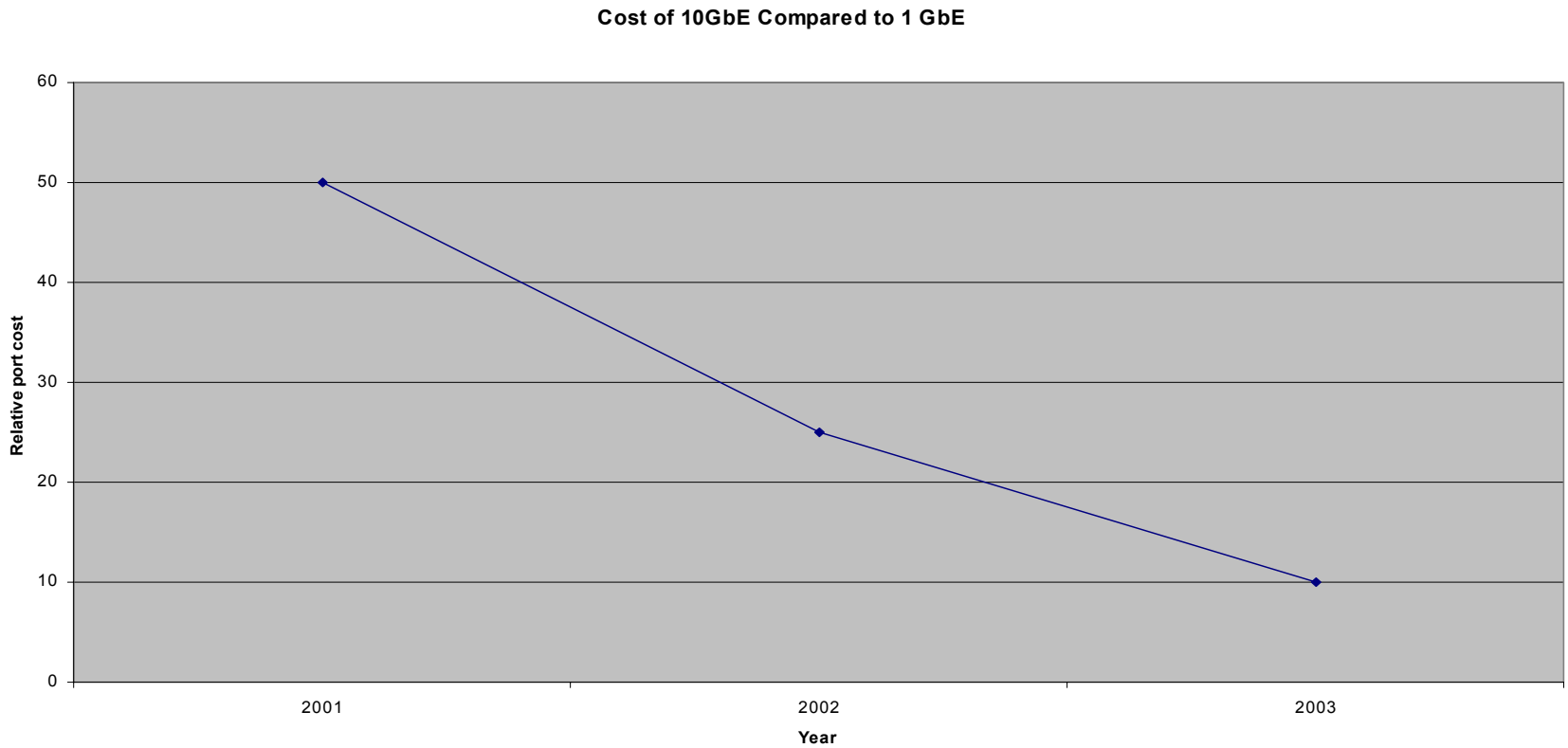
- **Currently 10 GbE has PMDs that support the following distances on optical fiber:**
  - **Up to 26 m on multimode (160 MHz*Km)**
    - **10GBASE-SR/SW (Table 52-6 802.3ae)**
  - **Up to 10 Km on single mode fiber**
    - **10GBASE-LR/LW (Table 52-11 802.3ae)**
  - **Up to 40 Km on single mode fiber**
    - **10GBASE-ER/EW (Table 52-15 802.3ae)**
- **No other parts shipping that I know of**
- **10GBASE-CX4 rapidly progressing**
  - **Will support a reach of up to 15 m on special copper cable**

Where does 10GBASE-T fit in the market?

- 850 nm PMD is just now shipping
- There is still no 10GBASE-LX4 PMD shipping
- Although customers can use multimode fiber for 10GbE designs, they do not save much compared to existing equipment
- That leaves 10GBASE-T to fill in the gap
- Why 10GBASE-T?

What are the drivers?

- **Simply put, 10 GbE still costs too much**
  - **Although the costs are coming down**



What are the drivers?

- **Machines today are capable of delivering roughly 50% of a 10 GbE in the best case**
  - **This means we spend a factor of 10 more to get at best, a 5 fold increase in throughput**
  - **Can't take full advantage of the BW**
  - **Need to scale with the cost of the network**
- **If we could get the cost down to a factor of three relative to GbE, then the customers actually save money by implementing systems that are simpler than the alternatives**

What are the alternatives?

- **Use Link Aggregation**
  - **4-port GbE interfaces just hitting the market**
  - **Have to install 5 x 2-port NICs (PCI) and use 10 switch/router ports**
  - **Or install 3 x 4-port GbE NIC (PCIX)**
    - **Leaves 2 unused ports and uses 10 switch/router ports**
  - **Configuration is not simple**
  - **Troubleshooting is even more complicated**
    - **How do you capture traffic for analysis since portions of the flow travel across different cables?**
    - **Could capture from each wire, but then you have to re-assemble the traffic to look at it chronologically**
  - **The bottom line is that can be done but its not easy**

Data Centers

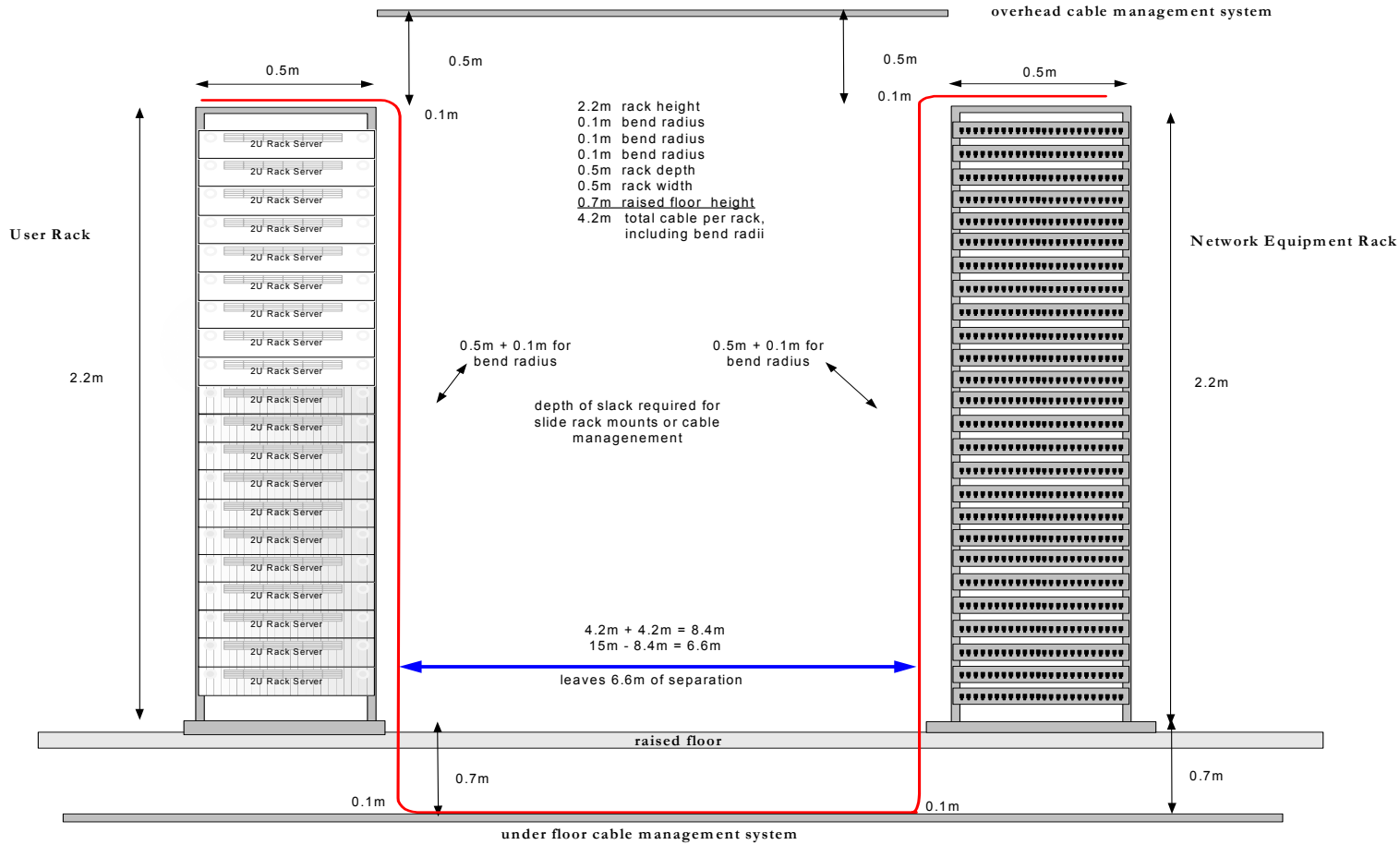
- High performance servers/clusters are the most likely place for 10 GbE in the near future
- These servers are typically deployed in data centers
- 10GBASE-T has been seen as the solution to deployment of 10 GbE in the data center
  - What problems are being solved?
    - Optical equipment costs too high
    - 10GBASE-CX4 doesn't cover the distance
- What are the characteristics of the data center?

Data Centers

- **Vary in size**
- **Cable distribution**
  - **Overhead cable trays**
  - **Under raised floor**
  - **Rack-to-rack cable management systems**
  - **Cables will not be “direct runs” meaning longer routes between racks**
- **What is the maximum distance between racks?**
  - **The cable must reach the entire height, width, and depth of the rack**
  - **There cannot be tight bends or kinks in the cable**
  - **By the time you add up the additional lengths due to the dimensions of the rack and access to the cable management system it adds 4 to 5m per rack**

Data Centers

- **4.2 m of cable in each rack (8.4m) allows only 6.6m between racks for 10GBASE-CX4 (15m)**



Data Centers

- There are four data centers at LBNL
  - LBLnet equipment room
  - IT services equipment room
  - NERSC computer room
  - ESnet network services equipment room
- All use star wired topology
  - Network equipment cabinet is located as close to center of the room as possible
  - All cabling is distributed from the network equipment cabinets
  - All cabling is category 5e (based on TIA-568 B.2)

Data Centers

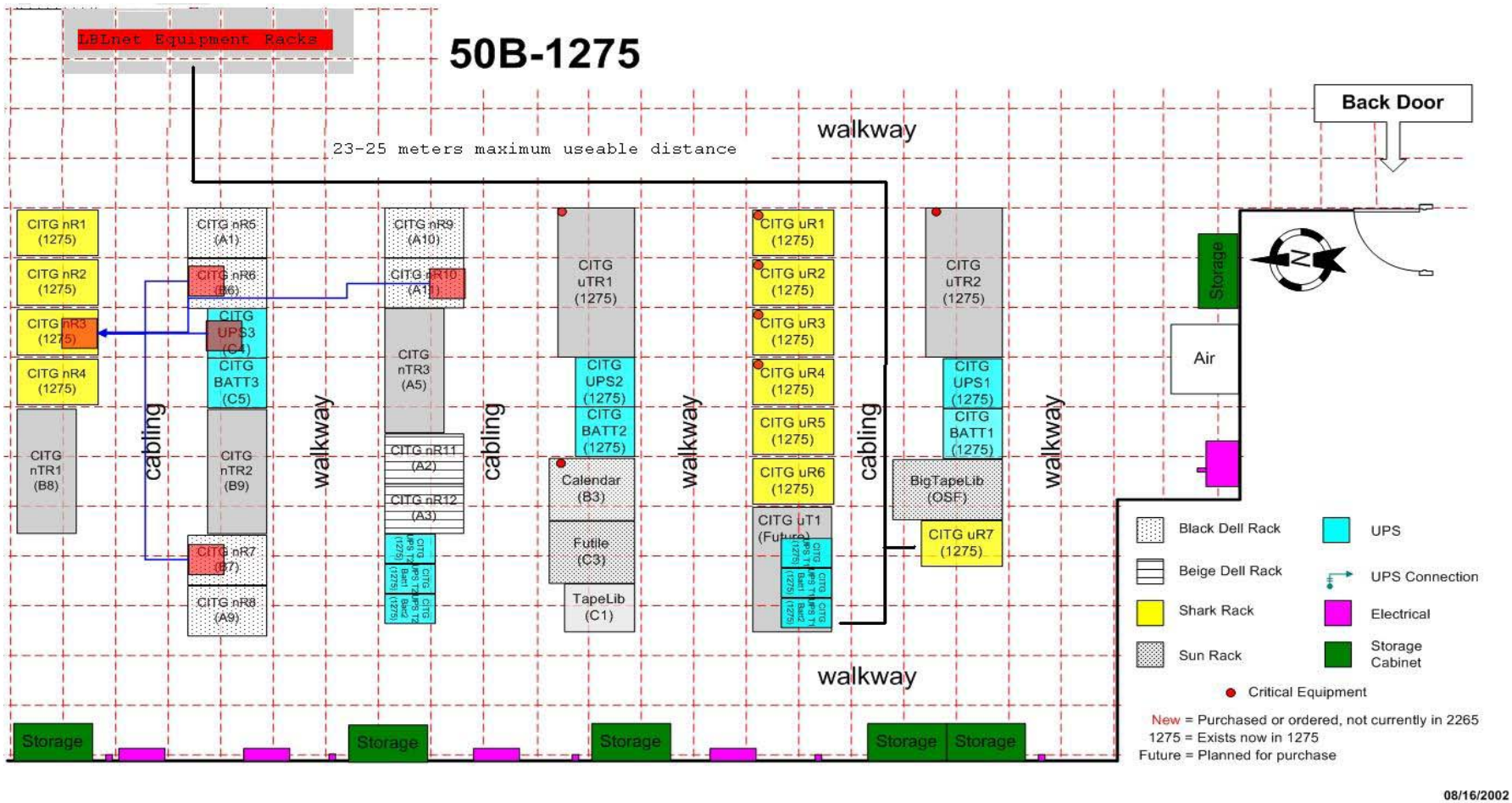
- **Examples:**
  - **LBLnet equipment room**
    - **325 square meters (3500 sq ft.)**
    - **Raised floor and overhead cable trays**
    - **32 racks**
    - **Distance between network equip. rack and farthest user rack: 27m**
  - **IT services equipment room**
    - **465 square meters (5000 sq. ft)**
    - **Raised floor and overhead cable trays**
    - **113 racks**
    - **Distance between network equip. rack and farthest user rack: 32m**

Data Centers

- **Examples:**
  - **NERSC computer room**
    - **1394 square meters (15,000 sq. feet)**
    - **Overhead cable trays and raised floor**
    - **250 racks**
    - **Distance between network equip. rack and farthest user rack: 48m**
  - **ESnet equipment room**
    - **465 square meters (5000 sq. ft)**
    - **Overhead cable trays**
    - **60 racks**
    - **Distance between network equip. rack and farthest user rack: 33m**

Data Centers

- **Here's a section of the IT services room:**
- **23-25m of horizontal structured cable**

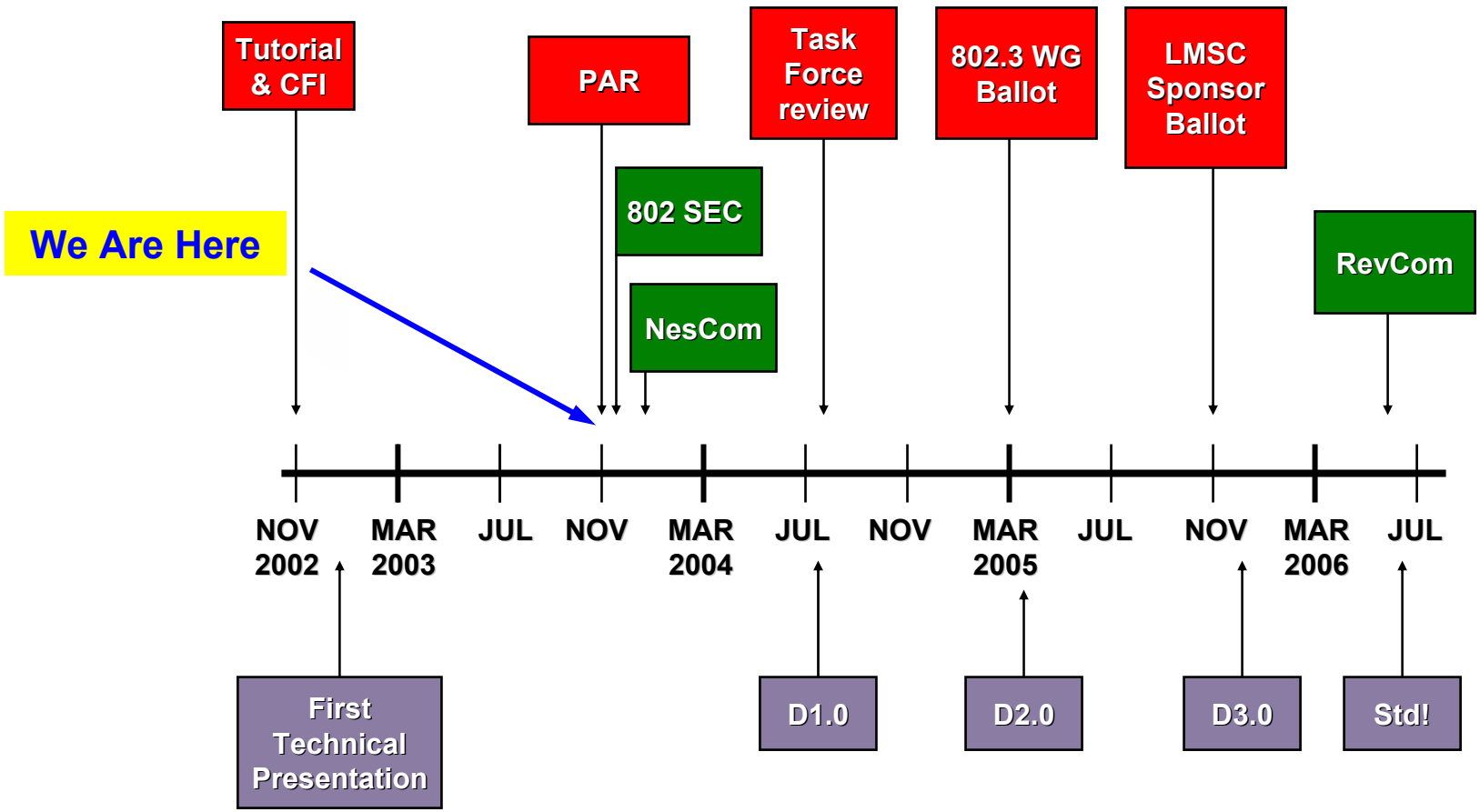


Data Centers

- **Notice that the maximum distance between network and user racks is under 50 meters on the average**
  - **Greater than 15 meters exceeds the distance covered by 10GBASE-CX4**
    - **Recall that the maximum distance between racks is 6.6m**
    - **There is a significant gap in the coverage**
  - **This is where 10GBASE-T is needed**

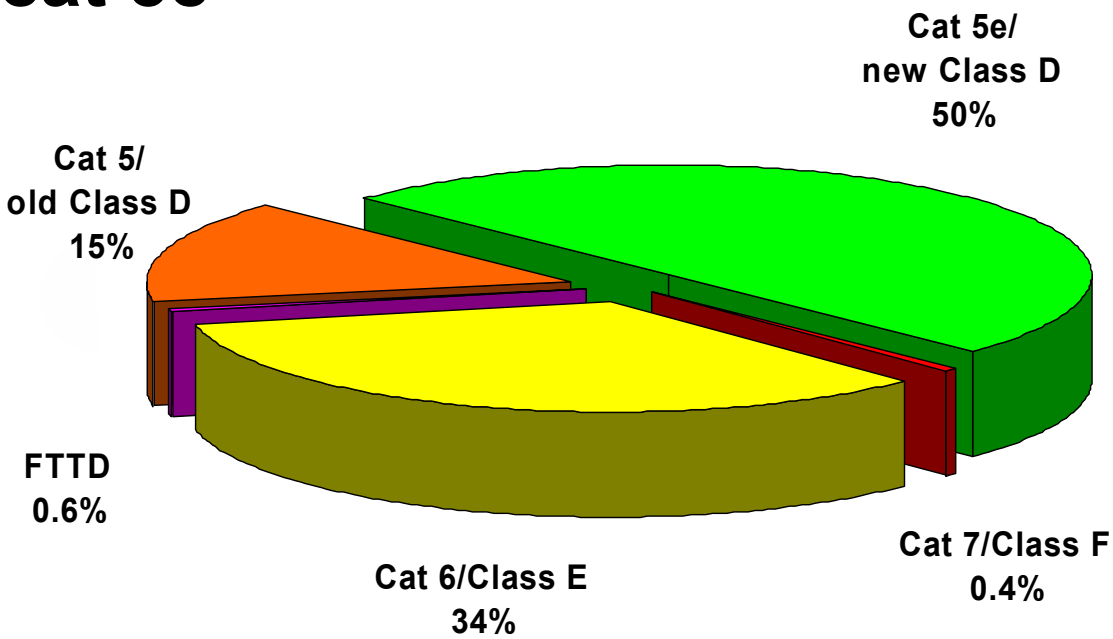
Questions

- **What about category 5e cable?**
  - **Assuming that 802.3 decides to approve the PAR**



Questions

- **When 10GBASE-T goes to sponsor ballot, a large percentage of the installed base will be cat 5e**

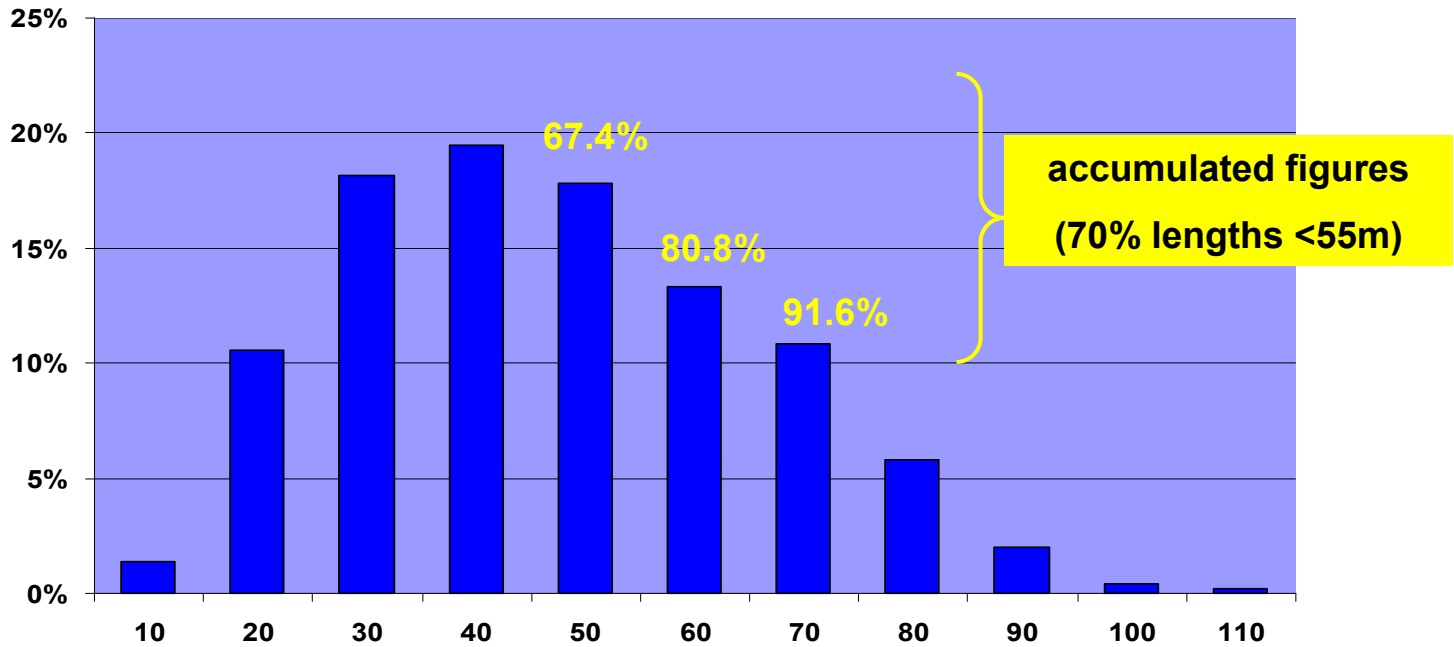


Source: Flatman_1_0303.pdf

Questions

- **More than half of the installed base of cables are less than 50m**

Source: Flatman, Fluke, Hubbell, Nordx/CDT, Siemon (total sample >108,000 cabling links)



Questions

- According to the presentation “10GBASE-T: How much capacity is enough?” the minimum channel capacity required to support 10GBASE-T is 18-20 Gb/s (roth_1_0503.pdf)
- According to the presentation “10GBASE-T Cabling Baseline Proposal” a 4 connector 50m category 5e channel has a capacity of 17.7 Gb/s with no alien crosstalk mitigation and 20.6 Gb/s with 5 dB alien crosstalk mitigation (diminico_2_0703.pdf)

Questions

- **Why not have an objective to specify some distance on category 5e?**
  - **50m is adequate coverage for data centers**
  - **It has been noted that some of the installed base of category 5e cables will not meet an extended frequency limit**
    - **Do we know what percentage fails?**
    - **As long as the channel parameters are specified, why not give users the opportunity to test the cables? If they fail, then we'll install new ones**
- **How will we be able to claim 10GBT has Broad Market Potential when there is no objective to support at least some distance on category 5e cable?**
  - **Better to answer this question now than in 802.3**

Conclusions

- The data centers are where the majority of high end computers are going to be deployed
  - This is where 10 GbE demand will most likely be
- Its clear that 10GBASE-T fills the gap for 10 Gigabit Ethernet needs in data centers
  - 10GBASE-CX4 will only support a 6.6m inter-rack distance when equipment is placed at the top of each rack
- 10GBASE-T has a much broader market potential with an objective to support the installed base of category 5e cable
  - This would meet the needs of LBNL and other DOE national facilities
  - Category 5e will still be a large percentage of installed cable when the standard is completed in 2006