

Surveying the Industry For a Higher Speed Ethernet

**Presented by
Mike Bennett, LBNL**

**802.3 Working Group Interim
Knoxville, TN**

Background

- **CFI “marketing team” prepared and sent surveys to poll users in industry**
 - **Good idea to find out what your potential customers actual need**
 - **Try to learn where the “pain” is and something about the installed cable infrastructure**
- **Survey Questions**
 - **Titled “Bandwidth / Cabling Survey: IEEE HSSG”**
 - **Cabling and Distance section categorized by application space**
 - **Service Provider Transport**
 - **Data Centers (Service Provider or enterprise)**
 - **Enterprise LAN**
 - **Note to R&E / Government organizations to fill-in as it applied and likely to be a mix of the others**

Background

- **Survey Questions (cont.)**
 - Titled “**Bandwidth / Cabling Survey: IEEE HSSG**”
 - **Bandwidth Trends section is more general (paraphrased):**
 - **If you have bottlenecks, where do they occur?**
 - **How do you expect BW/traffic needs to change over the next 5 years**
 - **How many bottlenecks do you have?**
 - **How many do you expect in 5 years?**
 - **What method do you use to scale physical links?**
 - **LAG(802.3ad) or ECMP?**
 - **How many ports in a “typical” LAG (802.3ad)**
 - **Do you have a cap on the number of ports in a LAG (802.3ad)**

Survey Respondents Summary

Organization	Business Model	Bandwidth Drivers
Yahoo!	Portal / Content	Broadband / Internet
Comcast	Residential Broadband	On Demand, HDTV
Cox Cable	Residential Broadband	P2P, Video
Cisco Systems	Network Vendor	DC Consolidation / Virtualization
AMS-IX	Internet Exchange	Broadband / Internet 10 GbE services Data center to data center links
Equinix	Internet Exchange	
LINX	Internet Exchange	
IX in Japan	Internet Exchange	
Level(3) Communications	Long Haul, ISP	
ESnet	ISP/Metro	
Brookhaven National Lab	Research	Computing Moving large data sets over LAN and WAN links
Fermi National Lab	Research	
Lawrence Berkeley Lab	Research	
Lawrence Livermore Lab	Research	
NERSC	Research	

Summary of Bandwidth and Growth Projections

Consumer Broadband Access



Broadband

Cox: 10 GbE today, BW growth 50-75% per year for next 3 – 5 years

Cisco: 10GbE today, 40 GbE needed in 5 years, 100 GbE preferred

Comcast: 3X BW increase in 3 to 5 years

Level 3: 8x10 GbE LAG today, BW growth 15x in 5 years (~70%/year)

Yahoo!: 4x10 GbE LAG today, BW doubling in <12 months

LLNL: 4x10 GbE LAG and 500x10 GbE ports today, 10x speed requirement in 5 years on deployed ports

Internet Backbone Networks

Internet Exchanges: Up to 8x10 GbE LAG today, BW growth 50-75% per year for next 3 – 5 years

ESnet: 10 GbE today, 10 Gbps on 20+ links 5 years from now; 5-10 locations will require more than 40 Gbps

Corporate Data Centers and Enterprise

Research Networks

and Government Facilities

Issues with 10 GbE LAG(802.3ad)

- **LAG increases cost**
 - LAG ports cannot be used for revenue as a customer port
 - Higher TCO for multiple fibers in Metro/WAN deployments
- **LAG increases complexity**
 - Difficult to plan for capacity and traffic engineering, especially in large networks (LAGs connected to LAGs)
 - Manageability/troubleshooting of multiple physical links for a single logical interface
- **Uneven distribution of traffic over LAGs**
 - Limitations and complexity in vendor implementations
 - Inefficient distributions of large flows

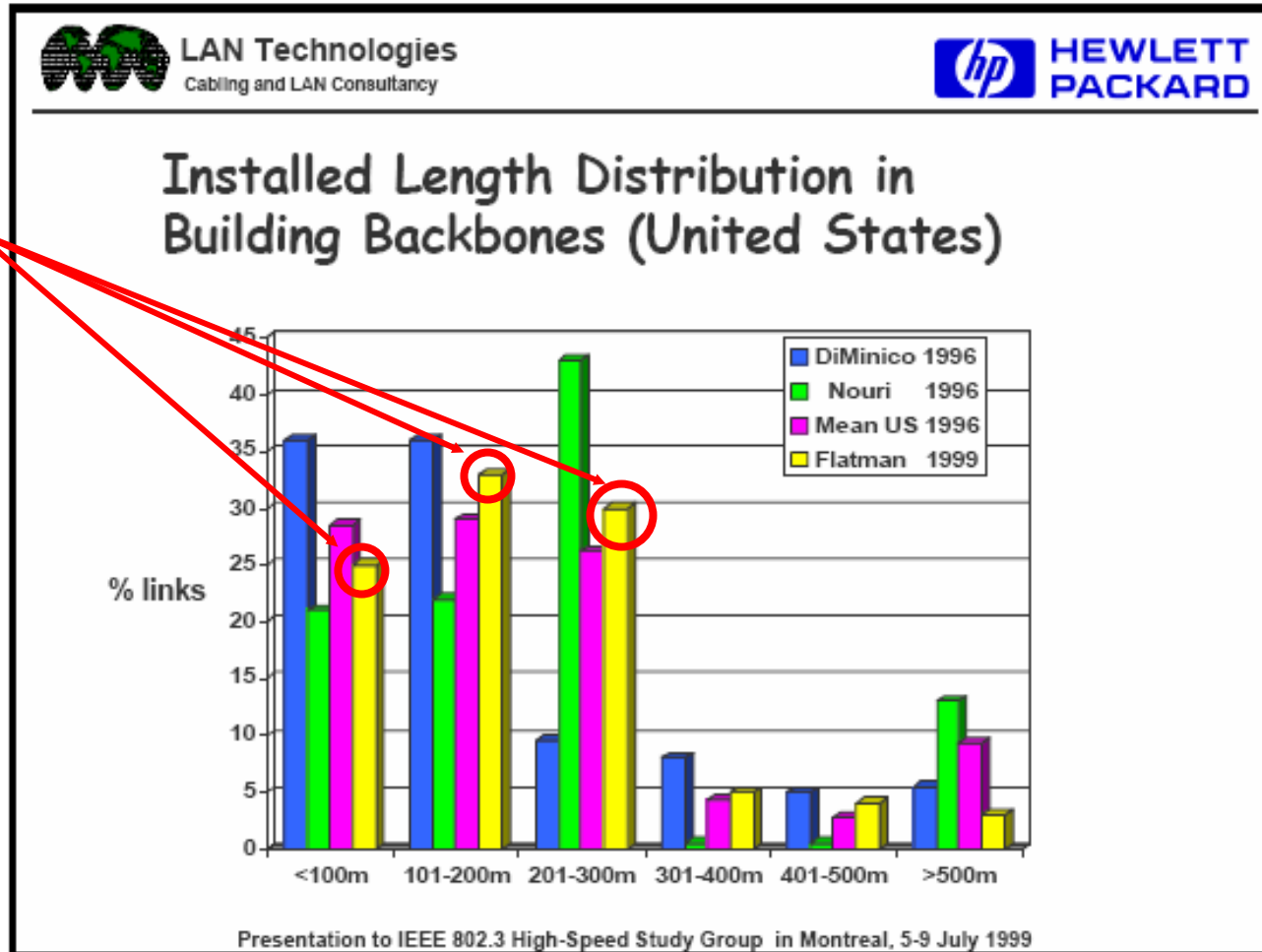
Typical Cable Distances in Meters

Data Center (10 Respondents)	Server to Access	Access to Aggregation	Aggregation to Core	Data Center to Data Center
Average Distance	70.3	121	582	15.7k
Median	70	70	30	2k
Range	3 to 200	20 to 400	10 to 2k	150 to 100k
Service Provider (9 Respondents)	Rack to Rack	Intra POP	Metro Access	
Average Distance	41.4	7.3k	59k	
Median	30	1.5k	60k	
Range	3 to 100	10 to 40k	1.5k to 100k	

60% of respondents identified applications < 300m

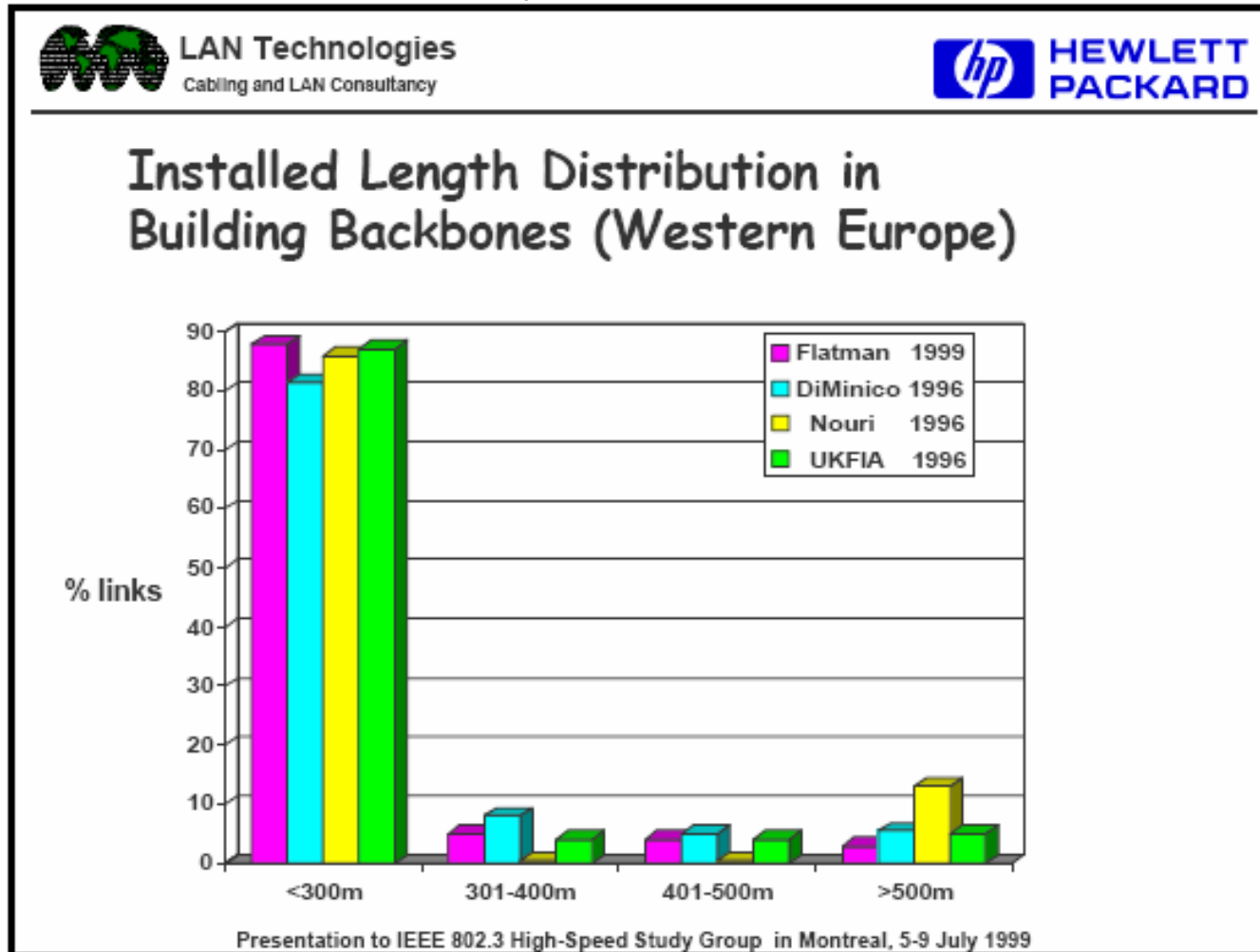
Typical Cable Distances in Meters

- Alan Flatman's survey data from 7/99 as a reference



Typical Cable Distances in Meters

- Alan Flatman's survey data from 7/99 as a reference



Interesting Comments

- **“I expect we need bandwidth 100Gbps or over 100Gbps with aggregated links next 5 years ”**
- **“I expect bw needs to double every year for the next 3 years, and then taper off after that”**
- **“We will aggregate with 8 links before 100GbE interface is released ”**
- **“Currently we have LAGs up to 8x10G, traffic projections point to needs of 64x-128x10G in the 2010 timeframe ”**
- **“In the (HPC) data center: Seeing much more 10GigE on servers now; majority 10GigE in 5 years. Will 100GigE arrive soon enough? ”**

Interesting Comments

- **“Already 10GE is not enough for customers to connect with and customers connect with aggregated 10GE connections”**
- **“We expect to be limited by the capabilities of the switches (8x LAG is typical max today). In addition we concerned about the effectiveness of flow based load balancing for larger link bundles particularly as flow sizes grow ”**
- **“Today we only use 2 members in a LAG due to lack of a standard method for sharing traffic between LAG members and equipment practices that make it costly to place all members of a group onto different line interface cards ”**
- **“For LAG’s we see more unpredictability during removal and insertion, we prefer to go to the next higher speed once we’ve hit 4”**

Conclusions

- **“LAN Reach”**
 - **Applications could include:**
 - **Datacenters, IntraPOP, InterPOP, Central Offices, Enterprise Backbone, Aggregation, Clustering/HPC, Storage, Building Risers, Bandwidth Exchange Links, etc.**
 - **Past standards - 100m, 220m, 300m and 2km**
 - **60% of respondents identified applications < 300m for HSSG links**
 - **Potential implementations could support 2km or more**
- **“WAN Reach”**
 - **Applications could include:**
 - **Inter-Office Facilities, Metro Networks, Datacenter Backup, Private Networks, Backhaul, Distribution/Collector Networks, etc.**
 - **Past standards- 40km, 80km, and more**
 - **Distances up to 100km could be desirable**
 - **Potential implementations could support 40km or more**

Conclusions

- **We may have to conduct another cable survey similar to the one Alan Flatman did in 1999 for 802.3ae**
 - **We should probably know what types of cables are the installed base now**
 - **I'll bet there's not much 160 MHz*Km MMF left**
 - **I'd be interested to know how much OM3 has been installed**
- **LAG(802.3ad) of 10 GbE links is problematic today and inadequate for the future**

Acknowledgments

- **Thanks to the following people for their contributions**
 - **Adam Bechtel, Yahoo!**
 - **Vince Bonafede, Brookhaven National Laboratory**
 - **Eli Dart, Energy Sciences Network (ESnet)**
 - **Phil Demar, Fermi National Laboratory**
 - **Brent Draney, National Energy Research Supercomputing Center (NERSC)**
 - **Alan Flatman, LAN Technologies**
 - **Steve Garrison, Force10 Networks**
 - **Prashant Ghandi, Cisco Systems**
 - **Greg Hankins, Force10 Networks**
 - **Mike Hughes, London Internet Exchange**
 - **Randy Kinsey, Cox Communications**
 - **Joe Lawrence, Level 3 Communications**
 - **Lane Patterson, Equinix**
 - **Ted Sopher Jr., Lawrence Berkeley National Lab**
 - **Henk Steenman, Amsterdam Internet Exchange**
 - **Dave Wiltzius, Lawrence Livermore National Lab**
 - **And those who wished to remain nameless (you know who you are!)**