40G SMF Ad-Hoc Summary

Alessandro Barbieri
abarbier@cisco.com

Supporters (1/2)

- David Cunningham, Avago
- Andreas Bechtolsheim, SUN
- George Young, ATT
- Martin Carroll, VZ
- Mark Nowell, Cisco
- Gary Nicholl, Cisco
- Scott Kipp, Brocade
- Steve Song, Excelight
- Petar Pepeljugoski, IBM
- Daryl McCarty, FedEX
- Pete Anslow, Nortel
- Justin Abbott, Gennum
- Norbert Folkens, JDSU
- Ghani Abbas, Ericsson
- Howard Frazier, Broadcom
- Matt Traverso, Opnext
- Jeffery J. Maki, Juniper Networks
- Alan Flatman
- Jim Cataudella, Bank of America
- Brad Turner, Juniper Networks
- David Ofelt, Juniper Networks
- Stephen Trowbridge, Alcatel-Lucent
- John Jaeger, Infinera
- Frank Chang, Vitesse
- Andy Moorwood, Infinera
- Drew Perkins, Infinera
- Mike Ressl, Hitachi Cable

Supporters (2/2)

- Jeffry Pereira , Schlumberger
- Michael C Bouchet, JPMorgan Chase
- Chris Deemer, State of Michigan
- Stephen Strong, Juniper Networks
- Terrance Rosadiuk, Alcatel-Lucent
- Serbay Murat , MergeOptics
- Fiere Julien , MergeOptics
- Chris Cole, Finisar
- Rachel Kartch, EDMC
- Hideki Isono (Fujitsu Ltd.)
- Kazuyuki Mori (Fujitsu Ltd)
- Atsushi Takai (Opnext)
- Ralf-Peter Braun, T-Systems
- Arne Alping, Ericsson
- Shimon Muller, SUN
- Jesse Simasarian, Alacatel-Lucent
- Osamu Ishida, NTT
- Shoukei Kobayashi, NTT
- Larry Green, IXIA
- Hiroshi Nakano (NEC Electronics Corp.)
- Ryoko Yoshimura (NEL)
- Shinji Nishimura (Hitachi)
- Hidehiro Toyoda (Hitachi)
- Rick Rabinovich, T-Packets
- Ali Ghiasi, Broadcom
- Hideaki Horikawa (Oki Electric Industry Co., Ltd.)

Agenda

- Material presented during the Ad-Hoc meetings
  - Broad Market Potential: the enterprise and the carrier views
  - Technical Feasibility
  - Economic Feasibility
- PAR and 5 Criteria
- Closing recommendation

40GbE SMF PMD Ad-hoc presentations recap

- **4 ad-hoc teleconferences were held where ~20-30 attendees per call confirmed attendance with an email to the chair.**
- **A total of 8 presentations (excluding closing and opening presentations) were reviewed around the topics of Technical Feasibility, Economic Feasibility and Broad Market Potential:**

  TF:

  http://www.ieee802.org/3/ba/public/AdHoc/40GSMF/Tsumura_40_01_0208.pdf

  http://www.ieee802.org/3/ba/public/AdHoc/40GSMF/traverso_40_01_0208.pdf

  http://www.ieee802.org/3/ba/public/AdHoc/40GSMF/cole_40_01_0208.pdf

  EF:

  http://www.ieee802.org/3/ba/public/AdHoc/40GSMF/cole_40_02_0208.pdf

  http://www.ieee802.org/3/ba/public/AdHoc/40GSMF/traverso_40_01_0308.pdf

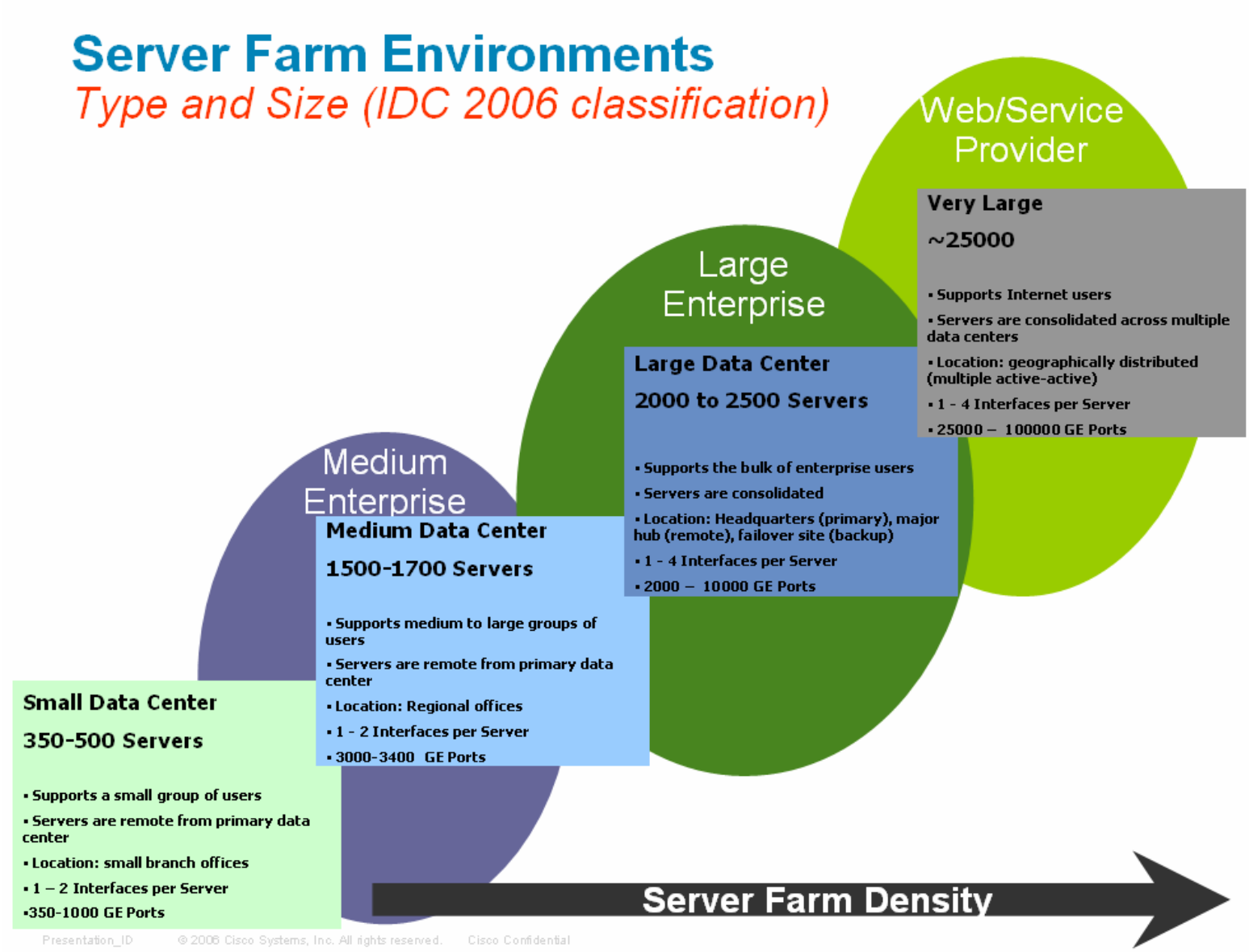
  BMP:

  http://www.ieee802.org/3/ba/public/AdHoc/40GSMF/carter_40_01_0208.pdf

  http://www.ieee802.org/3/ba/public/AdHoc/40GSMF/simsarian_40_01_0308.pdf

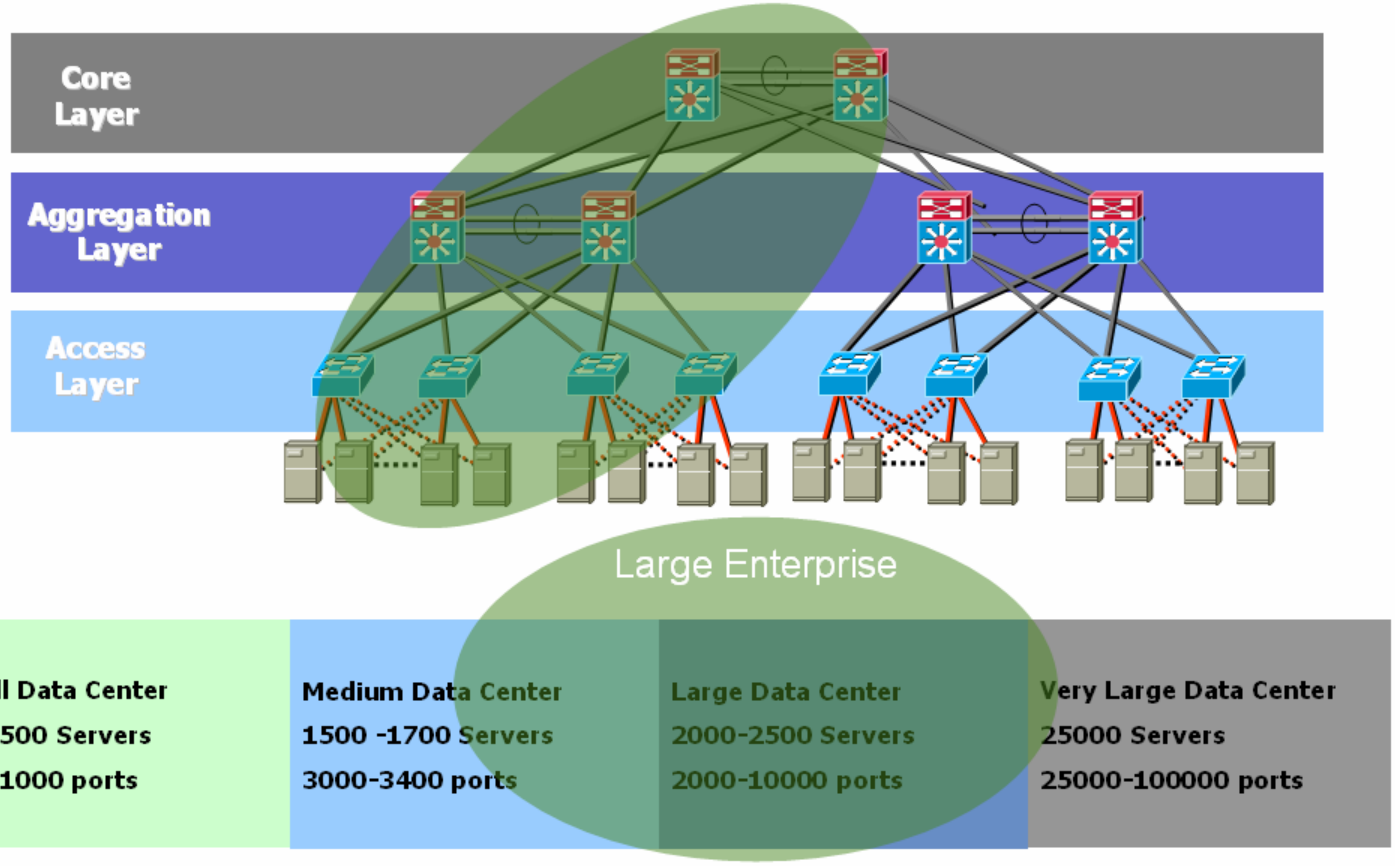
.

Different Classes of Data Centers
carter_40_01_0208



Different aggregation link speeds requirements
carter_40_01_0208

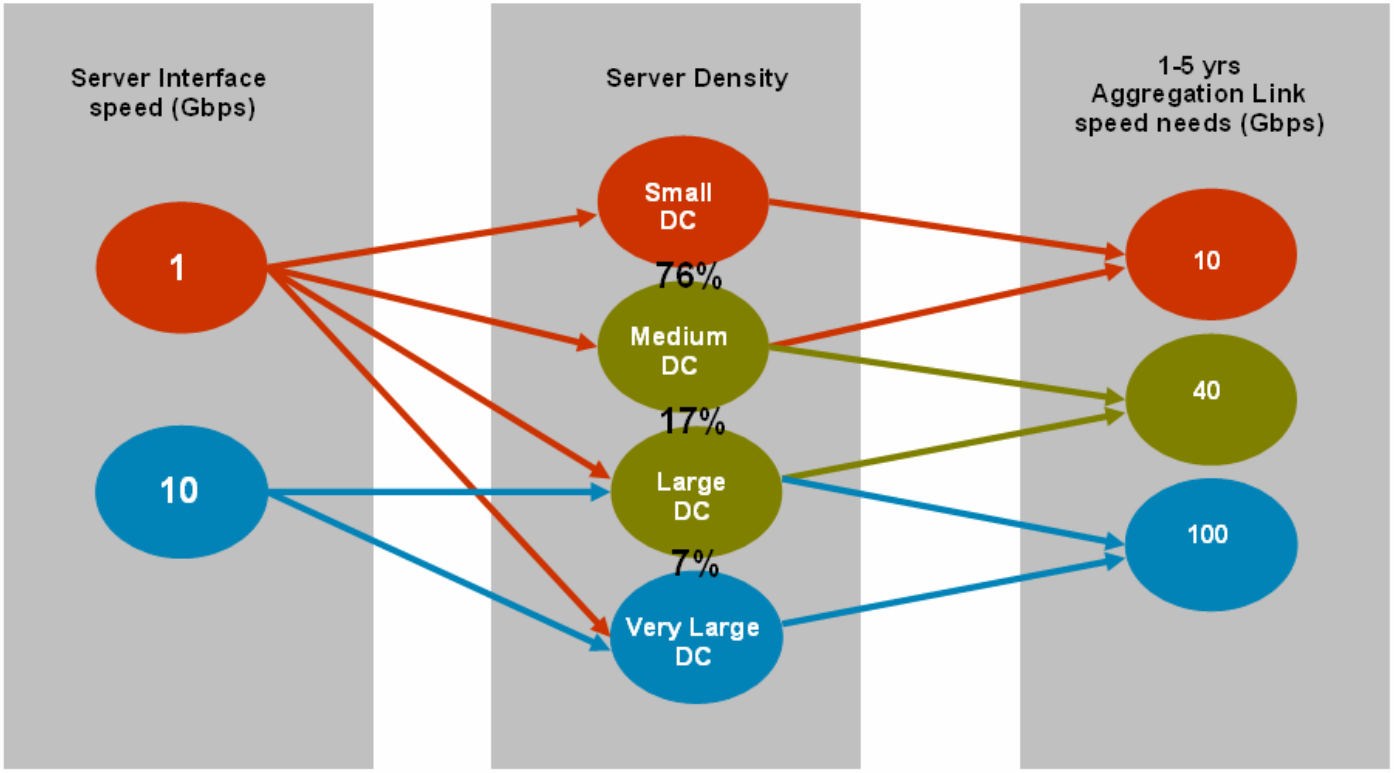
Server density and I/O speed determine depth and scale of the Data Center architecture layers



Tier 2 - 17% (IDC)

Different aggregation link speeds requirements
carter_40_01_0208

Switch aggregation speed needs in the Data Center



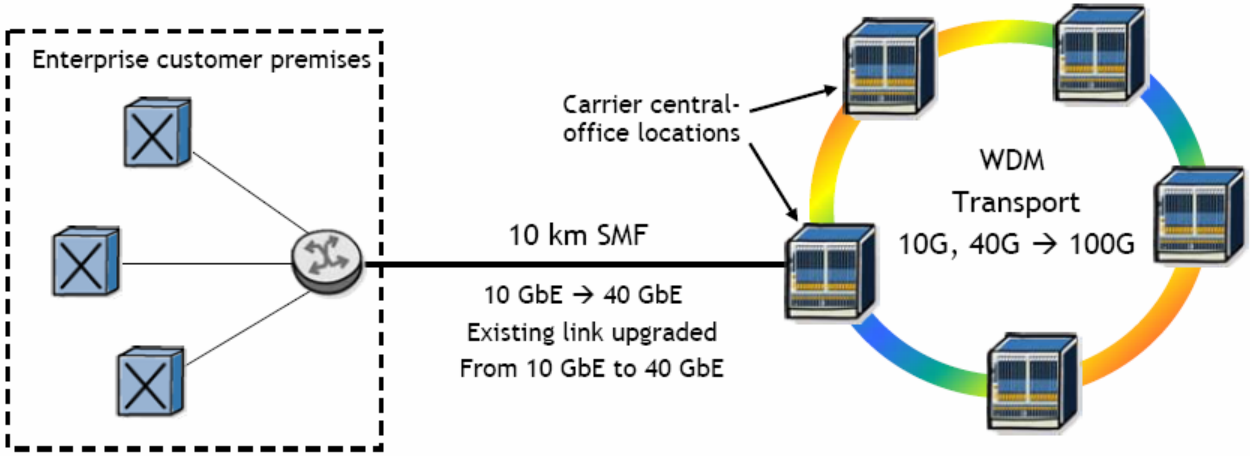
Different classes of Data Centers have different needs (and cost sensitivity) in terms of aggregation link speed.

40GbE SMF Broad Market Potential
The Enterprise view

- Different classes of Data Centers have different needs in terms of aggregation link speed determined by different requirements for server speed and server density.
- There is a broad market for 40GbE SMF to address the needs of Tier 2 and Tier 3 data center core links.

Carrier applications for 40G and 100G
simsarian_40_01_0308

40 GbE and 100 GbE Applications



- There will be a market for 100 GbE SMF driven by high-end applications - e.g. IPTV video on demand
- Web surfing: bandwidth per user ~ 100 kb/s vs. HD TV: bandwidth per MPEG4 HD stream ~ 8 Mb/s
- Example metropolitan area network with 1 million households and 10% concurrency:
  - Video on demand generates ~800 Gb/s
  - Web surfing generates ~10 Gb/s
- Data applications can often be cost-effectively supported by 40 GbE switch interfaces
  - Connection from enterprise data center to 40 Gb/s WAN transport
- Need for 100GbE and 40 GbE on single-mode fiber comes from the diversity of applications
  - High-end, bandwidth intensive (e.g. entertainment networking) vs. data centric (e.g. server traffic aggregation)

Alcatel·Lucent

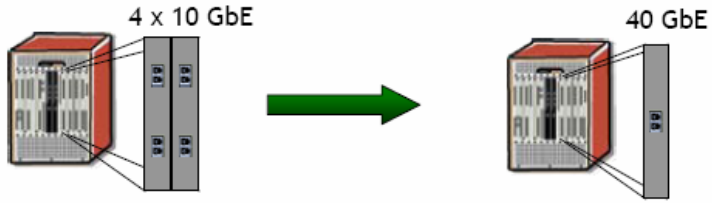
Inefficiencies of Link Aggregation in Carrier Networks
simsarian_40_01_0308

Problems with 4 x 10 GbE link aggregation

There are several reasons why link aggregation of 4 x 10 GbE is not satisfactory for many customers

- Fiber may not be available:

- Lower port density for multiple low-bitrate ports and increased OPEX from a larger number of managed ports



- Flow-based load balancing can be problematic (see trowbridge_01_0906.pdf)
  - E.g., encrypted flows with IPSEC

Why 10km for 40GbE SMF

- During a teleconference representatives from the Carrier community indicated that SMF is required to connect Enterprise customers to Central Office locations because the Enterprise and Carrier equipment is increasingly not co-located for Ethernet-based WAN interconnection to OTN/WDM transport services.
- 10km ensures adequate coverage from a Carrier perspective to reach a broad population of their Enterprise customers for shared WAN transport.
- 10km would be consistent with de-facto or standard reaches for GbE, 10GbE, 100GbE.

40GbE SMF Broad Market Potential
The Carrier view

- Certain applications (e.g. internet data) in carrier network can realize the benefits of 40GbE, while others bandwidth intensive applications (e.g. video) will demand 100G.
- 40GbE overcomes the inefficiencies of 4x10G.
- SMF support for 40GbE is consistent with the objective to support OTN.

Broad Market Potential Summary

- Both in the Enterprise and Carrier space separates sets of end users and applications with different bandwidth needs have been identified.
- Given the distinct application characteristics and customer profiles for 40GbE and 100GbE, the impact of 40GbE SMF on the volumes of 100GbE SMF is going to be marginal.
- As a result the overall Broad Market Potential of the IEEE 802.3ba project is augmented by adding support for SMF at 40GbE.

Technical Feasibility

- Three component companies have presented to the ad-hoc on a duplex SMF 10km interface.
- All three showed similar approach using a 4x10G CWDM multiplexed onto a single fiber. One also showed a 40G serial proposal.
- Adapting the 10GE link model for 40Gb, draft specifications have been proposed to achieve 10km specification with margin
- By utilizing off the shelf optical components and IC's, it can be said with a high confidence that the reliability should be similar to that of a 10Gbase-LX-4 module.

Technical Feasibility
tsumura_40_01_0208

ExceLight

Summary

- Feasibility of 40GE SMF 10km is investigated. CWDM grid DML is expected to be the most viable solution
- Both 1310nm band and 1550nm band CWDM are feasible. 1550nm band shows small advantage in optical power output
- Recommend more discussions on wavelength to consider other possible applications:
  - Longer distance support such as 25 or 40km reach on SMF
  - OM3 duplex 100m support with 4x10G CWDM using similar architecture as 10Gbase-LX4

Technical Feasibility
cole_40_01_0208

4x10G 1310nm DML 10km SMF Power Budget

10G Link Budget 10km SMF TP2 → TP3	CWDM DML λ = 1331nm* ER = 3.5dB
Fiber Loss (G.652 A&B)	4.3 dB
Penalties (10GBASE-L)	3.2 **
Connector loss	2.0
Total budget	9.5 dB

10G Pwr. Budget 10km SMF TP2 → TP3 OMA	CWDM DML λ = 1331nm* ER = 3.5dB
TX Min → Max	2.5→ 5.5 dBm
TP2 TX Min 2.5dB Mux loss	0.0 **
Link Budget (dB)	9.5 dB
TP3 RX Min 2.5dB DeMux loss	-9.5
RX Min (with crosstalk penalty)	-12.0 dBm

* Additional Link Budget and Power Budget analysis at λ = 1331nm and λ = 1271nm required to confirm final Budget numbers

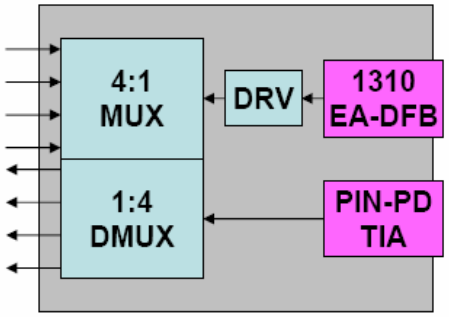
** TP2 TX Min can be reduced for lower cost by reducing Penalties using today's DML technology transmitter parameters (vs. 10GE parameters)

Finisar

Technical Feasibility
traverso_40_01_0208 ("Sped up LX4" vs. Serial)

40GbE SMF: Serial

opnext



Wavelength	1300 to 1324	nm
SMSR	35	dB
TX OMA	+2.5	dBm
TX Avg.	+0.73	dBm
ER	8.5	dB
RIN	-132	dB/Hz
RX OMA	-6.5	dBm

10km is *challenging* but achievable

0.48dB margin at 10km

Basics	Input=	**Bold**	Ts(20-80)	**10** ps
	Q=	**7.04**	Ts(10-90)	15 ps
	Base Rate=	**41250** MBd	RIN(OMA)	**-132** dB/Hz
Transmitter			RIN at MinER	-134.5 dB/Hz
Wavelength Uc		**1300** nm	RIN_Coef=	**0.70**
Uw (see notes)		**0.10** nm	Det.Jitter	**1.0** ps inc.
Tx pwr OMA=		**2.50** dBm	DCD_DJ=	**1** ps TP3
Min. Ext Ratio=		8.50 dB	Effect. DJ=	0.00 (UI) ex
"Worst"ave.TxPwr		**0.73** dBm	MPN k(OMA)	**0**
Ext. ratio penalty		1.24 dBo	Tx eye height	62.7%
Tx mask X1=		**0.3** UI	Refl Tx	**-12** dB
X2=		**0.4** UI	ModalNoisePen	**0** dB
Y1=		**0.25**	Tx mask top	0.2 UI

PolMD DGDmax = 7.3ps per ITU recommendation for 0.3UI for DGD

5

Economic Feasibility

- Cost factors for Ethernet components and systems are well known.
- Economic feasibility for 40GbE is understood is HSSG, need to understand whether there is any change by adding one more SMF PMD objective.
- Two optical component manufacturers have presented that the cost of a 10km SMF component will be between 5-10x cost of a 10GBase–LR module.
- Network design, installation and maintenance costs are minimized by preserving, network architecture, management and software

Economic Feasibility
cole_40_02_0208

40GE 10km 1310nm PMD Economic Feasibility

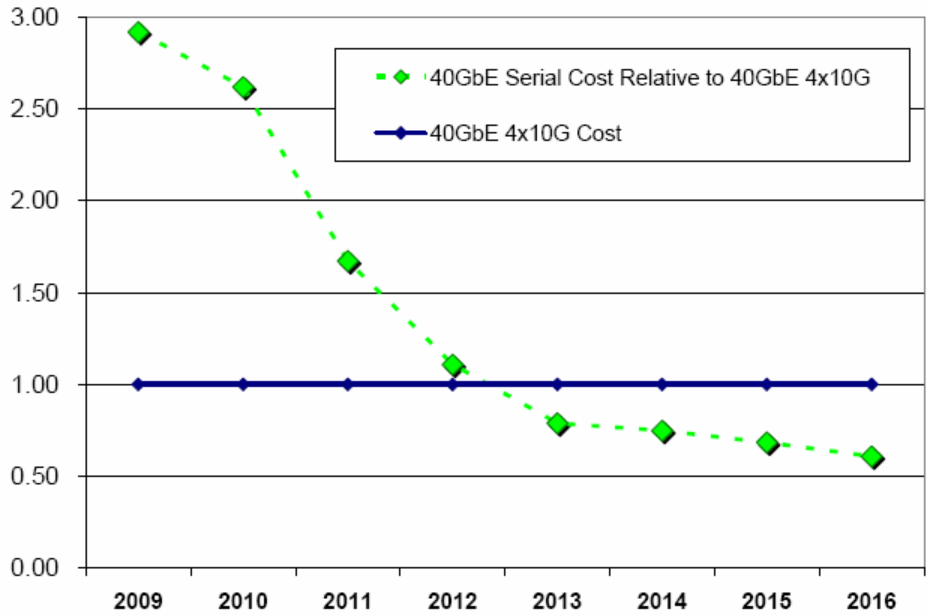
- 40GE 10km SMF 4x10G 1310nm un-cooled DML PMD is Economically Feasible.
- Economic Feasibility is enhanced (development cost driven) by leveraging existing 10GE-LR Transmitter, 10GE-LRM Receiver, and 100GE form factor technology.
- Economic Feasibility is enhanced (yield driven) by taking advantage of new 10G DML technology and relaxing 802.3 10GE-LR Transmitter assumptions to reduce the 10km Link Budget and improve DML yield.
- Economic Feasibility is enhanced (volume driven) by using 40GE SMF PMD for 100m duplex OM3 MMF applications.
- Economic Feasibility is enhanced (volume driven) by 40GE addressing near term cost sensitivity markets that can not be addressed by 100GE

  Gen1 100GE SMF PMD cost (p6) is 8x to 12x Gen1 40GE 10km SMF PMD cost (p4)

Economic Feasibility
traverso_40_01_0308 ("Sped up LX4" vs. Serial)

40GbE SMF: Serial vs. Parallel (LX40) Relative Cost

opnext



- "Same" volumes considered
- Volume of 40GbE assumed to be "small" in 2009 thru 2010
- As 40GHz interfaces become feasible in CMOS, the 40GbE will drop in cost significantly
- Reuse of 10G elements is not practical to achieve significant relative cost reduction

- **In SMF, parallel has been shown to be less cost competitive than serial**

Agenda

- Material presented during the Ad-Hoc meetings
  - Broad Market Potential: the enterprise and the carrier views
  - Technical Feasibility
  - Economic Feasibility
- PAR and 5 Criteria
- Closing Recommendation

PAR and 5 Criteria

- Current 802.3ba PAR is unaffected by inclusion of a 40GbE SMF objective.
- There is a requirement at the EC that the 5 Criteria responses need to stay aligned with a TF's objectives.
- For the TF to approve a motion to add an objective, it needs to know that the 5 criteria can be modified adequately.
- Modifications have been proposed and will need to be word-smithed and approved by TF if the objective is first approved by the TF.

Broad Market Potential (1 of 2)

- **Broad sets of applications**
- **Multiple vendors and numerous users**
- **Balanced cost (LAN versus attached stations)**

- **Bandwidth requirements for computing and core networking applications are growing at different rates, which necessitates the definition of two distinct data rates for the next generation of Ethernet networks in order to address these applications:**
  - **Servers, high performance computing clusters, blade servers, storage area networks and network attached storage all currently make use of 1G and 10G Ethernet, with significant growth of 10G projected in ’07 and ’08. I/O bandwidth projections for server and computing applications, including server traffic aggregation, indicate that there will be a significant market potential for a 40 Gb/s Ethernet interface.**
  - **Core networking applications have demonstrated the need for bandwidth beyond existing capabilities and the projected bandwidth requirements for computing applications. Switching, routing, and aggregation in data centers, internet exchanges and service provider peering points, and high bandwidth applications, such as video on demand and high performance computing environments, have demonstrated the need for a 100 Gb/s Ethernet interface.**

Proposed modification

Broad Market Potential (2 of 2) [NO MODIFICATIONS]

- **Broad sets of applications**
- **Multiple vendors and numerous users**
- **Balanced cost (LAN versus attached stations)**

- **There has been wide attendance and participation in the study group by end users, equipment manufacturers and component suppliers. It is anticipated that there will be sufficient participation to effectively complete the standardization process.**
- **Prior experience scaling IEEE 802.3 and contributions to the study group indicates:**
  - **40 Gb/s Ethernet will provide approximately the same cost balance between the LAN and the attached stations as 10 Gb/s Ethernet.**
  - **The cost distribution between routers, switches, and the infrastructure remains acceptably balanced for 100 Gb/s Ethernet.**
- **Given the topologies of the networks and intended applications, early deployment will be driven by key aggregation & high-bandwidth interconnect points. This is unlike the higher volume end system application typical for 10/100/1000 Mb/s Ethernet, and as such, the initial volumes for 100 Gb/s Ethernet are anticipated to be more modest than the lower speeds. This does not imply a reduction in the need or value of 100 Gb/s Ethernet to address the stated applications.**

Compatibility [NO MODIFICATIONS]

- **IEEE 802 defines a family of standards. All standards shall be in conformance with the IEEE 802.1 Architecture, Management, and Interworking documents as follows: 802. Overview and Architecture, 802.1D, 802.1Q, and parts of 802.1f. If any variances in conformance emerge, they shall be thoroughly disclosed and reviewed with 802. Each standard in the IEEE 802 family of standards shall include a definition of managed objects that are compatible with systems management standards.**

- **As an amendment to IEEE Std 802.3, the proposed project will remain in conformance with the IEEE 802 Overview and Architecture as well as the bridging standards IEEE Std 802.1D and IEEE Std 802.1Q.**
- **As an amendment to IEEE Std 802.3, the proposed project will follow the existing format and structure of IEEE 802.3 MIB definitions providing a protocol independent specification of managed objects (IEEE Std 802.1F).**
- **The proposed amendment will conform to the full-duplex operating mode of the IEEE 802.3 MAC.**
- **As was the case in previous IEEE 802.3 amendments, new physical layers specific to either 40 Gb/s or 100 Gb/s operation will be defined.**
- **By utilizing the existing IEEE 802.3 MAC protocol, this proposed amendment will maintain maximum compatibility with the installed base of Ethernet nodes.**

Distinct Identity [NO MODIFICATIONS]

- **Substantially different from other IEEE 802 standards**
- **One unique solution per problem (not two solutions to a problem)**
- **Easy for the document reader to select the relevant specification**

- **The proposed amendment is an upgrade path for IEEE 802.3 users, based on the IEEE 802.3 MAC.**
- **The established benefits of the IEEE 802.3 MAC include:**
  - Deterministic, highly efficient full-duplex operation mode
  - Well-characterized and understood operating behavior
  - Broad base of expertise in suppliers and customers
  - Straightforward bridging between networks at different data rates
- **The Management Information Base (MIB) for IEEE 802.3 will be extended in a manner consistent with the IEEE 802.3 MIB for 10 / 100 / 1000 / 10000 Mb/s operation.**
- **The proposed amendment to the existing IEEE 802.3 standard will be formatted as a collection of new clauses, making it easy for the reader to select the relevant specification.**
- **Bandwidth requirements for computing and networking applications are growing at different rates. These applications have different cost / performance requirements, which necessitates two distinct data rates, 40 Gb/s and 100 Gb/s.**

Technical Feasibility

- **Demonstrated system feasibility**
- **Proven technology, reasonable testing**
- **Confidence in reliability**

- **The principle of scaling the IEEE 802.3 MAC to higher speeds has been well established by previous work within IEEE 802.3.**
- **The principle of building bridging equipment which performs rate adaptation between IEEE 802.3 networks operating at different speeds has been amply demonstrated by the broad set of product offerings that bridge between 10, 100, 1000, and 10000 Mb/s.**
- **Systems with an aggregate bandwidth of greater than or equal to 100 Gb/s have been demonstrated and deployed in operational networks.**
- **The proposed project will build on the array of Ethernet component and system design experience, and the broad knowledge base of Ethernet network operation.**

  The experience gained in the development and deployment of 10 Gb/s technology is applicable to the development of specifications for components at higher speeds. For example, parallel transmission techniques allow reuse of 10 Gb/s technology and testing.

  Component vendors have presented data on the feasibility of the necessary components for higher speed solutions. Proposals, which either leverage existing technologies or employ new technologies, have been provided.

- **The reliability of Ethernet components and systems can be projected in the target environments with a high degree of confidence. Presentations demonstrating this have been provided.**

Economic Feasibility

- **Known cost factors, reliable data**
- **Reasonable cost for performance**
- **Consideration of installation costs**

- **The cost factors for Ethernet components and systems are well known. The proposed project may introduce new cost factors which can be quantified.**
- **Presentations indicate that for the server market and computing applications, including server traffic aggregation, the optimized rate to provide the best balance of performance and cost is 40 Gb/s. For the network aggregation market and core networking applications, the optimized rate offering the best balance of performance and cost is 100 Gb/s.**
- **In consideration of installation costs, the project is expected to use proven and familiar media, including optical fiber, backplanes, and copper cabling technology.**
- **Network design, installation and maintenance costs are minimized by preserving network architecture, management, and software.**

Proposed modification

Ad-Hoc conclusion and recommendation to the Task Force

- **During the closing conference call the Ad-Hoc group (the chair reckoned roughly 50 people by counting online attendees on the web conference applications) expressed unanimous support to make a motion into the Task Force to add 10km SMF support as an objective for 40GbE.**
- **This is the recommended motion to be made later in the week during this TF meeting:**

  Move that the 802.3ba taskforce adopt the following objective in replacement of existing 40Gb/s Physical layer objective:

  Provide Physical Layer specifications which support 40 Gb/s operation over:

  at least 10km on SMF

  at least 100m on OM3 MMF

  at least 10m over a copper cable assembly

  at least 1m over a backplane