100 Gbit/s Ethernet for True End-to-End Carrier-Grade Ethernet Networks

R. H. Derksen, A. Kirstädter, G. Lehmann

Siemens AG, Corporate Technology, Information & Communications, Germany
Tel.: +49-89-636-43460, Fax: +48-89-636-45814, E-mail: rainer.derksen@siemens.com

Abstract: Ethernet is an unbeaten success story extending its reach from LAN to the metro areas and probably soon also into core networks. 100 Gbit/s Ethernet will be a key enabler for true end-to-end carrier-grade Ethernet networks. We discuss functions and standards required to enable carrier-grade Ethernet backbone networks. We report as well a CAPEX and OPEX study of Ethernet core networks and competitive network architectures as on 100 Gbit/s transmission experiments over several hundred kilometres of fibre using a cost-effective integrated ETDM receiver. The results propose that Ethernet will soon be mature enough for deployment in backbone networks and provide huge cost advantages to providers.

1. Introduction

Backbone networks represent the top of the carrier's network hierarchy. They connect networks of different cities, regions, countries, or continents. The complexity of these technologies imposes substantial financial burdens on network operators, both in the area of Capital Expenditures (CAPEX) and Operational Expenditures (OPEX).

The Ethernet protocol is a possible enabler of more cost-efficient backbone networks, as it is characterized by simplicity, flexibility, interoperability, and low costs. While Ethernet is traditionally a Local Area Network (LAN) technology, continuous developments already enabled its deployment in Metropolitan Area Networks (MANs). Recent research and standardization efforts aim at speeding up Ethernet to 100 Gbit/s, resolving scalability issues and supplying Ethernet with carrier-grade features. For this reason, Ethernet might in the near future become an attractive choice and serious competitor in the market of backbone networks.

The first part of this paper elaborates on requirements and possible architectures of carrier-grade Ethernet-based core networks. The requirements in the areas of Quality-of-Service (QoS), resilience, network management, and scalability for introducing carrier-grade features into Ethernet will be outlined. Different network architectures and related introduction strategies will also be explained briefly.

The second part of this paper examines the economics of Ethernet networks in comparison to SONET/SDH-based network architectures. Hands-on business cases propose that both CAPEX and OPEX show a superior performance for Ethernet.

In the third part we provide an overview on high-speed transmission experiments. We motivate the need of integrated direct 100 Gbit/s modulation and reception and give details of an integrated high-speed electrical receiver originally designed for 86 Gbit/s. A transmission experiment with this receiver at 100 Gbit/s is described.

2. Requirements for Carrier-Grade Ethernet-based Core Networks

Several alternatives are possible for the transfer of Ethernet frames over a transport network, e.g. a native Ethernet solution that uses MAC-in-MAC encapsulation [1] or the transfer via a MPLS-based backbone network [2].

Both architectures require a control plane to provide for network-wide functionality of signalling, traffic engineering, quality-of-service, and protection, established e.g. via Generalized-MPLS (GMPLS) [3]. Table 1 shows the mapping of protocol functionality to the requirements of backbone networks for both architectures mentioned above. Functionality that is provided or controlled by GMPLS is also included in the illustration. It is clearly visible that a lot of functional redundancies still exist in these stacks.

	provide high bandwidth	QoS Features	Resilience Mechanisms	OAM Features
IP		DiffServ	Rerouting	Ping etc.
1) native Ethernet		BW reservation (GMPLS) + Eth priorities	VLAN prot sw or RSVP-TE fast reroute	IEEE 802.1ag Connectivity Fault Mngt
2) MPLS Ethernet		BW reservation (GMPLS) + LSP priorities	LSP prot. sw or RSVP-TE fast reroute	Fault detection of nodes and paths
Ethernet PHY	100Gbps			
WDM	DWDM		standby channels	
Fiber	bundles		redundant fiber deployment	

Table 1: Protocol layer functionality in Ethernet core networks.

2.1 Carrier-Grade Functionality

In order to be suited for core networks Ethernet needs carrier gradeness, i.e. it has to offer and implement the following functionalities [4]:

End-to-End **Quality of Service** enabling service providers to guarantee and enforce transmission quality parameters (e.g. bandwidth, jitter, delay) according to a specified service-level agreement (SLA) and connection acceptance control with the customer. A corresponding QoS framework is currently developed by the Metro Ethernet Forum (MEF) [5]. QoS-conform forwarding in Ethernet switches will also likely be controlled via GMPLS.

Resilience mechanisms with the required scalability and convergence-time behaviour – beyond the current spanning tree bridging and its extensions. E.g. using GMPLS to manage protection of links and paths in carrier-grade Ethernet networks, pre-provision backup paths and switch over in the case of failure (c.f. the MEF protection framework [6]).

Enhanced **Operations Administration & Maintenance** (OAM) is indispensable in carrier networks, as it enables failure detection, localization, and performance monitoring (c.f. the IEEE 802.1ag standard extension or the OAM framework set up by the MEF [7]).

As Ethernet originates from the LAN, it faces several **scalability** issues on its progress into Wide Area Networks. Address space limitations will be resolved via the upcoming IEEE Provider Backbone Bridge standard [1] and similar approaches. The utilization of meshed network structures has to be enabled e.g., via extensions to the spanning tree protocol (STP) or hierarchies of VLAN structures. Apart from the maximum Ethernet frame size of 1500 Byte (extended via jumbo frames) the maximum transmission distance of Ethernet signals is another topic that has to be considered carefully upon the deployment of Ethernet in backbone networks. Here, node distances are usually in the area of a few hundred kilometres or even more - i.e. much larger than the maximum transmission distance of 10 G Ethernet (10GE) (70-80 km according to vendor specifications). However, the effort spent on extending the reach of Ethernet signals is rewarded by equipment savings. Figure 1 illustrates for a generic German backbone network the possible port count savings in an Ethernet core network where optical grooming can be applied up to the maximum transmission distance avoiding unnecessary electrical processing of transit traffic. A very promising approach in this direction is reported below in section 4.
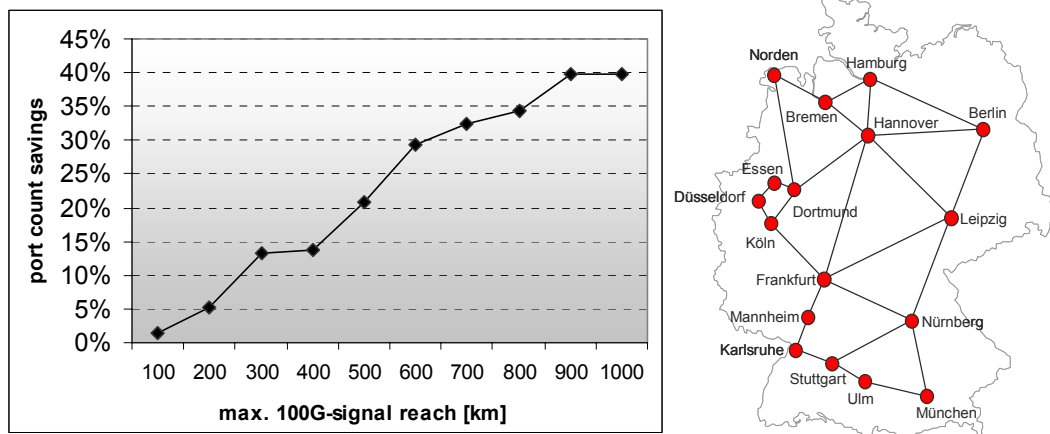

Figure 1: Port count savings in grooming-enabled Ethernet networks.

3.   CAPEX and OPEX Performance of Ethernet Core Networks

3.1 CAPEX: Modelling Approach and Results

In order to calculate the total CAPEX of a specific network architecture, future traffic loads, network device counts, and network device prices have to be estimated. The German reference network above was used for the physical topology of the considered backbone and corresponding traffic matrices were extrapolated to determine future link loads. The traffic is assumed to grow homogenously at a rate of 40% per annum. A shortest-path routing algorithm is applied to determine the single link loads from which the number of switches, routers, and line card ports were then obtained depending on the network architecture.

Notes: The cost of WDM-equipment and fibre were not included as they are independent of the network architecture choice. Further, The SONET/SDH business cases use an incremental CAPEX scenario (upgrading the existing network) whereas in the Ethernet business cases migration from SONET/SDH had to be respected.

Future equipment prices were extrapolated following a careful analysis of market data, past price developments, and price relations between Ethernet and SONET/SDH (for details on this approach see [4]).

3.2 Architectures and Results

The following generic network architectures were considered:

(a) IP/POS-over-WDM: SONET/SDH over WDM is used for node-to-node IP packet transfer between Label Edge Routers (LERs) and Label Switch Routers (LSRs). A 1+1 protection scheme is applied.

(b) IP/POS-over-SDH-over-WDM: Same as (a). The traffic is additionally switched and groomed along SDH add-drop-multiplexers and cross-connects inside the core.

(c) IP/POS-over-OXC-over-WDM: Same as (b), the SDH switches being replaced by optical cross-connects (OXCs), the optical signal range being large enough.

(d) IP/MPLS-over-Ethernet-over-WDM: LERs at the network edge and MPLS-enabled Ethernet switches in the core (Ethernet over MPLS, c.f. section 2). 1:1 protection is applied - i.e. all capacity is overprovisioned by 100%.

(e) Ethernet-over-WDM: The core being a native Ethernet network with Ethernet switches both at edge and core: Ethernet traffic does not have to traverse LERs at the ingress and egress points of the backbone. A few LERs handle the small share (30%) of traffic that requires IP routing. 1:1 protection is applied.

(f) Ethernet-over-WDM with service-level protection: Same as (f), only premium traffic (share 30%) is protected.
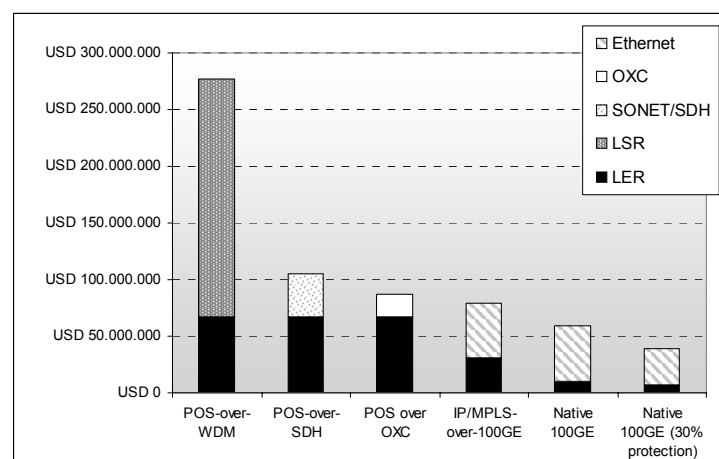

Figure 2: Accumulated CAPEX comparison

The CAPEX results (aggregated for 2009 to 2012) split up into components belonging to the different layers are illustrated in Figure 2. The high prices of POS interfaces generally lead to a high cost component of LERs for all SDH-related infrastructures. On the transport layer, a pure POS-over-WDM network creates the most costs as only expensive router interfaces are used. POS-over-SDH architectures prove to be much cheaper as the SDH network employs SDH switches and interfaces instead of expensive LSR equipment for core switching. A POS-over-OXC network has an even better CAPEX performance due to lower switch and optical transceiver prices of OXC hardware.

Without exception the Ethernet business cases perform better than SDH architectures. In the MPLS Ethernet business case, still a considerable amount of CAPEX is related to expensive LERs and interfaces. A native 100GbE network enables higher savings in the LER category. Via a service-level differentiated protection scheme the CAPEX can be reduced even further.

3.3 OPEX COMPARISON

While OPEX generally include a lot more categories, the repair process is selected here since the impact of 100GbE can be predicted most visibly in this area. *Note*: As WDM failures and fibre breaks again are not considered. Thus, the total OPEX repair process values are very low.

The OPEX were evaluated for each network architecture (and year) of section 3.2 via first determining the total number and type of equipment. By using availability figures [8] the average repair time for a given backbone architecture was estimated. The related costs were derived by multiplying the total repair time with the average salary of a field or point-of-presence technician. Table 2 shows the results of the repair process OPEX, accumulated over the years 2009 to 2012 (OPEX rising with network growth):

NW architecture	Accumulated cost [USD]
(a) POS over WDM	515.000
(b) POS over SDH	340.000
(c) POS over OXC	490.000
(d) IP/MPLS over 100GE	245.000
(e) Native 100GE	245.000
(f) Native 100GE, 30% protected	170.000

Table 2: OPEX repair process comparison.

Due to the reduced device count (less switches and line cards) enabled by 100GbE instead of 40G POS Ethernet networks are more economical than SDH-architectures. A service-level protection scheme may further reduce overprovisioning and the related OPEX.

Further OPEX savings may be enabled via the provisioning of Ethernet services in comparison to legacy services like e.g. leased line or Frame Relay (study conducted by the MEF [9]. Thus, Ethernet architectures enable considerable savings in a wide range of OPEX areas.

4. 100 Gbit/s Transmission Experiments Using an Integrated ETDM Receiver

4.1 High-speed transmission: From OTDM to ETDM

Around year 2000 many activities have been started to increase the data rate per wavelength to 160 Gbit/s. Since no electrical processing was (and still is) possible at this high data rate optical processing has been required. The data stream was optically multiplexed and demultiplexed in time, the so called OTDM technique. A huge effort has been made to find methods for ultra-fast switches to extract bits out of a 160 Gbit/s OTDM data stream as by using four-wave-mixing [10][11][12], electro-absorption modulators [13], cross phase modulation [14][15], and Mach-Zehnder interferometer with integrated SOAs [16]. The functionality of time-domain add-drop multiplexers has even been proven with simultaneous transmission over 275 km field deployed fibre [17].

The transmission of high speed data rates above 100 Gbit/s is well understood and can be managed. As a consequence, the knowledge to realise the transmission of a 100GbE signal is present. The problem to solve is to find efficient electro-optical and opto-electrical conversion techniques. The above mentioned OTDM techniques are still

too complex and difficult to implement in commercial products. Therefore, electrical solutions are preferable to handle the data at the transmitter and receiver.

Recently, transmission experiments with direct modulation of 107 Gbit/s data signal [18] and direct detection and electrical demultiplexing of a 100 Gbit/s data signal [19] have been shown. The ability to process the data electronically at these data rates is a big step towards product integration.

4.2 Integrated ETDM receiver

By using ultra fast electronic circuits instead of elaborate optical methods in high-capacity optical transmission systems cost per transmitted bit per second and kilometre can be reduced. Electronic circuitry for 40 Gbit/s is already commercially available. A receiver for only 85 Gbit/s based on electrical time division multiplexing (ETDM) has been reported already in [20]. It consisted of a number of small single circuits and received only one of eight 10.6-Gbit/s channels at a time; furthermore, it required an additional photo diode for the clock recovery. To really exploit the cost advantage of an electrical receiver compared to optical solutions a compact integrated device is needed (preferably a single chip). As an important step towards 100 Gbit/s Ethernet we presented an integrated ETDM receiver, which comprised 1:2-demultiplexing (DEMUX) and clock & data recovery (CDR) on a single chip. The ETDM receiver was tested in a 100 Gbit/s transmission experiment and error-free performance after 480 km dispersion managed fibre (DMF) was achieved [19]. This was the first demonstration of a single chip ETDM receiver at 100 Gbit/s in a transmission experiment.
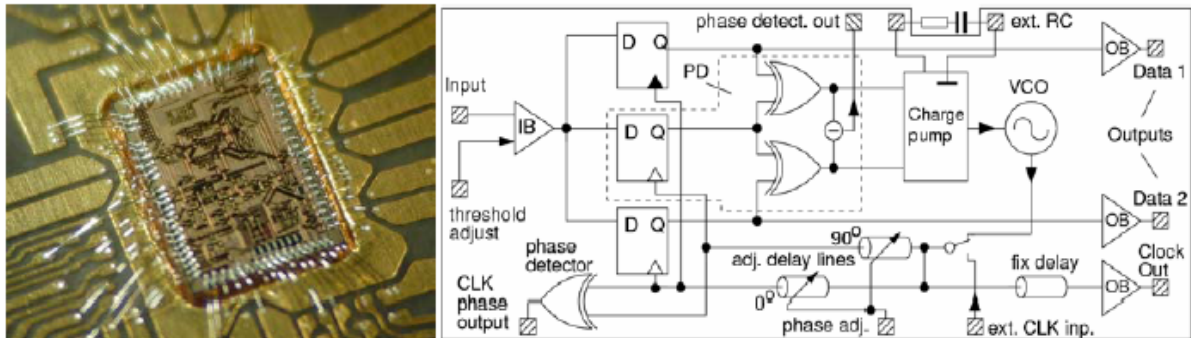

Figure 3: Photo (left) and block diagram (right) of the integrated ETDM receiver chip.

A block diagram of the integrated ETDM receiver chip is shown in Figure 3. The ETDM receiver chip was designed for 80 Gbit/s operation with forward error correction (FEC, 86 Gbit/s). It comprises the 1:2-DEMUX, the CDR and an integrated 43 GHz voltage controlled oscillator (VCO) and was successfully operated in an 86 Gbit/s transmission. Despite the fact that the chip was intended for 86 Gbit/s the innovative design enabled operation at more than 100 Gbit/s. The internal VCO was not designed for such an extremely large tuning range, so we had to use an external 50 GHz VCO together with an external control circuit driven by the chip-internally generated phase-detector signal. However, the integration of a 50 GHz VCO with the semiconductor process used for the receiver chip would be no problem [21].

The receiver chip was fabricated with a SiGe-bipolar transistor process (180 GHz transit frequency, 200 GHz maximum oscillation frequency) of Infineon Technologies. A detailed description of the circuit including its electrical characterisation is given in [22]. The 1:2-DEMUX can be operated up to 100 Gbit/s at the input, resulting in 2 x 50 Gbit/s at the output. A minimum voltage swing of 120 mV (single ended) is needed at the

input. The voltage swing at the outputs into an external 50 Ohm load is 250 mV single ended and 500 mV differential. The CDR provides recovered 50 GHz and 25 GHz clock output signals. The whole receiver chip requires a single supply voltage of -5 V and has a power dissipation of 5 W. All inputs and outputs are internally terminated with 50 Ohm. The receiver chip includes about 1000 transistors and has a size of 1.7 mm x 2.5 mm.

Transmission experiment

A schematic of the experimental setup comprising the 100 Gbit/s optical time division multiplexing (OTDM) transmitter, the 480 km fibre link and the 100 Gbit/s ETDM receiver is shown in Figure 4. Details can be found in [19], here we confine ourselves to some results
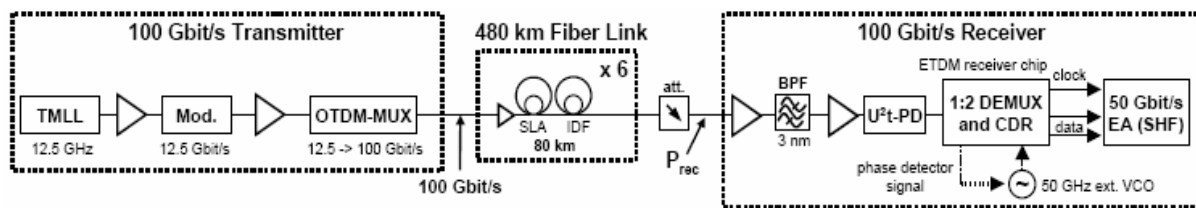

Figure 4: Schematic depiction of the transmission experimental setup.

The experimental results are shown in Figure 5. On the left hand side the bit-error ratio (BER) measured for the 100 Gbit/s data signal is shown as a function of the optical power P_{rec}. Error-free performance (BER=10^{-9}) was achieved back-to-back and after transmission over 480 km of dispersion managed fibre for a word length of 508 bits. The receiver chip was also tested back-to-back at a longer word length of 2^{31}-1 and a penalty of 2 dB was found, which can be partly attributed to a signal degradation in the transmitter.

Both demultiplexed 50 Gbit/s tributaries were measured and the results were identical. Eye diagrams of both 50 Gbit/s tributaries are shown on the right hand side of Figure 5, together with eye diagrams of the 100 Gbit/s data signal back-to-back and after transmission at the input of the receiver. The eye diagrams at 100 Gbit/s were measured with a high-bandwidth photodiode and a fast electrical sampling oscilloscope (70 GHz bandwidth). The required optical signal-to-noise ratio (OSNR) for a BER of 10^{-9} was 28.6 and 30.6 dB back-to-back and after transmission, respectively.

In the meantime, with the availability of the adequate measurement equipment, we also demonstrated error-free transmission over the same distance of fibre for 107 Gbit/s [23].
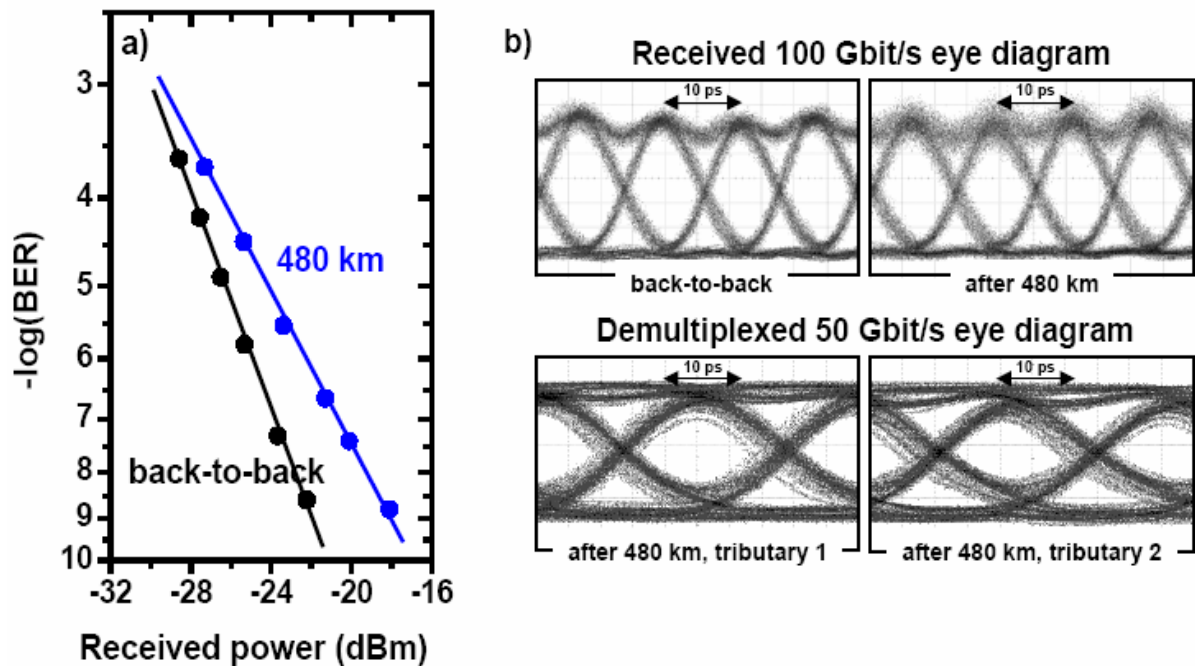
Figure 5: a) Bit error measurements back-to-back and after transmission over 480 km dispersion managed fibre,
b) top: 100 Gbit/s eye diagrams at the receiver back-to-back and after 480 km dispersion managed fibre, bottom: 50 Gbit/s eye diagrams after the electrical demultiplexer for both tributaries.

4.3 ETDM transmitter

Pure electrical multiplexing up to 100 Gbit/s is already feasible, e. g. [22], but the modulators require at these bit rates driving voltages which are difficult to achieve. Nevertheless, we are also working on this subject and first solutions are in sight – so that possibly within a short time transmission experiments with fully electrical transmitter and receiver can be accomplished.

Direct modulation with a bit rate of 107 Gbit/s was shown recently in a transmission experiment [18], but still needing the combination with an adaptive optical filter to equalize the distorted electrical signal before transmitting.

5. Conclusion

Optical data networks became the backbone of the data networks – and simultaneously the backbone of the communication and information society. The development of this network has been driven by and optimized for the telephone. To maintain the effectiveness of this network an alteration towards a data packet optimised network is required. The decision which technique will be the first choice for this step will be driven by the cost of the suggested solutions, CAPEX as well as OPEX. At the same time, the quality of the network has to keep the high standard provided today.

Ethernet evolved from LAN into MAN covering speeds up to 10 Gbit/s. A careful analysis of the required protocol features like network resilience, QoS, and OAM shows many redundancies within the layers of today's network architectures that have to be

resolved shaping a new end-to-end Ethernet layer with the required scalability and providing the required "Carrier-gradeness".

A CAPEX and OPEX analysis demonstrates a considerable cost advantage of 100GbE in comparison to SDH-based solutions. The superior CAPEX performance results from a huge cost advantage of Ethernet devices and their fast price decline. The reduced switch and line card count in 100GbE networks and the efficient economics of Ethernet services are responsible for a superior OPEX performance.

Recently, a lot of progress has been made towards cost effective solutions for such high data rates, as has been shown in several transmission experiments. The next step will be to realise real product solutions out of the prototypes which perfectly support the new network architecture.

Ethernet has a promising future in core networks, not just as link technology supporting an upper routing layer, but as a complete, cost-effective, and service-oriented infrastructure layer in the area of core networks. The industry-wide efforts to cover remaining challenges also confirm this outlook.

Acknowledgements

References

[1] IEEE, "Provider Backbone Bridges," [Online]. Available: http://www.ieee802.org/1/pages/802.1ah.html

[2] IETF, "MPLS Label Stack Encoding," (RFC 3032) [Online]. Available: http://www.ietf.org/rfc/rfc3032.txt

[3] D. Papadimitriou, J. Choi, "A Framework for Generalized MPLS (GMPLS) Ethernet," IETF Internet Draft [Online]. Available:http://www.ietf.org/internet-drafts/draft-papadimitriou-ccamp-gmpls-ethernet-framework-00.txt

[4] A. Schmid-Egger, A. Kirstädter, "Ethernet in Core Networks:: A Technical and Economical Analysis", Proc. HPSR2006, Poznan, Poland, June 2006.

[5] S. Khandekar, "Developing A QoS Framework For Metro Ethernet Networks To Effectively Support Carrier-Class SLAs," [Online]. Available: http://www.metroethernetforum.org/Presentations/IIR_DevelopingQoSFramework.pdf

[6] MEF, "Requirements and Framework for Ethernet Service Protection in Metro Ethernet Networks", [Online]. Available: http://www.metroethernetforum.org/PDFs/Standards/MEF2.pdf

[7] M. Squire, "Metro Ethernet Forum OAM," [Online]. Available: http://www.metroethernetforum.org/presentations/MEF%20OAM%202003-12-01.pdf

[8] EU-IST NOBEL, "Availability Model and OPEX related parameters", to be published .

[9] MEF, "Service Provider Study: Operating Expenditures," [Online]. Available for MEF members: http://www.metroethernetforum.org

[10] S.L. Jansen, M. Heid, S. Spaelter, E. Meissner, C.-J.Weiske, A. Schoepflin, G.D. Khoe, H. de Waardt," Demultiplexing a 160 Gbit/s OTDM signal to 40 Gbit/s by FWM in a SOA", Electronics Letters, Vol. 38 (17), Aug 2002 pp. 978 –980

[11] T. Yamamoto, E. Yoshida, and M. Nakazawa, "Ultrafast nonlinear optical loop mirror for demultiplexing 640 Gbit/s TDM signals", Electron. Lett. 34, 1013 (1998)

[12] T. Morioka, H. Takara, S. Kawanishi, K. Uchiyama, M. Saruwatari, "Polarisation-independent all-optical demultiplexing up to 200Gbit/s using four-wave mixing in a semiconductor laser amplifier" , Electron. Lett. 32, 840 (1996)

[13] B. Mikkelsen, G. Raybon, R.J. Essiambre "160Gb/s TDM transmission systems", Proceedings of ECOC'2000, paper 6.1.1

[14] M. Nakazawa, T. Yamamoto, K.R. Tamura, "1.28 Tbit/s–70 km OTDM transmission using third- and fourth-order simultaneous dispersion compensation with a phase modulator", Electron. Lett. 36, 2027 (2000)

[15] Feiste, R. Ludwig, C. Schubert, J. Berger, C. Schmidt, H.G. Weber, B. Schmauss, A. Munk, B. Buchold, D. Briggmann, F. Kueppers, F. Rumpf, "160 Gbit/s Transmission over 116 km Field-Installed Fiber Using 160 Gbit/s OTDM and 40 Gbit/s ETDM", Proceedings of OFC'2001, paper ThF3-1

[16] M. Heid, S. Spälter, G. Mohs, A. Färbert, W. Vogt, H. Melchior; "160-Gbit/s demultiplexing based on a monolithically integrated Mach-Zehnder interferometer", ECOC 2001 PD.B.1.8

[17] J.P. Turkiewicz, E. Tangdiongga, G. Lehmann, H. Rohde, W. Schairer, Y.R. Zhou, E.S.R. Sikora, A. Lord, D.B. Payne, G.D. Khoe, and H. de Waardt, "160 Gb/s OTDM Networking using Deployed Fiber", J. Lightwave Technol. (invited), Volume 23, Number 1, January 2005, Pages:225-235

[18] G. Raybon, P. J. Winzer, C. R. Doerr, "10 x 107-Gbit/s Electronically Multiplexed and Optically Equalized NRZ Transmission over 400 km, Proc. OFC 2006, PD paper PDP32, Anaheim, California (U. S. A.), March 2006.

[19] R. H. Derksen, G. Lehmann, C.-J. Weiske, C. Schubert, R. Ludwig, S. Ferber, C. Schmidt-Langhorst, M. Möller, J. Lutz, "Integrated 100 Gbit/s ETDM Receiver in a Tranmission Experment over 480 km DMF", Proc. OFC 2006, PD paper PDP37, Anaheim, California (U. S. A.), March 2006.

[20] K. Schuh, B. Junginger, H. Rempp, P. Klose, D. Rösener, E. Lach, "85.4 Gbit/s ETDM Transmission over 401 km SSMF Applying UFEC", *Proc. ECOC 2005*, PD paper Th 4.1.4, pp. 7-8, Glasgow (Unidted Kingdom), Sept. 2005.

[21] H. Li, H.-M. Rein, A. Schwerd, "SiGe VCOs operating up to 88 GHz, also suited for automotive radar sensors", Electronics Letters, vol. 39, pp. 1326-1327, Sept. 2003.

[22] U. Dumler, M. Möller, A. Bielik, T. Ellermeyer,, H. Langenhagen, W. Walthes, J. Mejri, "86 Gbit/s receiver module with high sensitivity for 160 x 86 Gbit/s DWDM system", Electronics Letters, vol. 42, no. 1, pp. 28-29, Jan. 2006.

[23] C. Schubert, R. H. Derksen, M. Möller, R. Ludwig, C.-J. Weiske, J. Lutz, S. Ferber, C. Schmidt-Langhorst, "107 Gbit/s Transmission Using An Integrated ETDM Receiver", submitted to ECOC 2006, Cannes, (France), Sept. 2006.