

Carrier-Grade Ethernet for Packet Core Networks

A. Kirstädter*^a, C. Gruber^a, J. Riedl^a, T. Bauschert^b

^aSiemens AG, Corporate Technology, Information & Communications;

^bSiemens AG, Communications

ABSTRACT

Ethernet is a permanent success story, extending its reach from LAN and metro areas now also into core networks. 100 Gbit/s Ethernet will be the key enabler for a new generation of true end-to-end carrier grade Ethernet networks. This paper first focuses on functionality and standards required to enable carrier-grade Ethernet-based core networks and possible Ethernet backbone network architectures will be discussed. The second part then evaluates the CAPEX and OPEX performance of Ethernet core networks and competitive network architectures. The results propose that Ethernet will not only soon be mature enough for deployment in backbone networks but also provide huge cost advantages to providers. A novel complete, cost-effective and service-oriented infrastructure layer in the area of core networks will arise. The industry-wide efforts to cover remaining challenges also confirm this outlook.

Keywords: Core networks, Ethernet, carrier-grade networks.

1. INTRODUCTION

Backbone networks represent the top of the carriers' 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 cost. 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 next section of this paper starts with describing the requirements and possible architectures of carrier-grade Ethernet-based backbone networks. The concepts in the areas of Quality-of-Service (QoS), resilience, and network management for introducing carrier-grade features into Ethernet will be outlined. Network and protocol layer architectures suited to introduce high degrees of scalability to the network will be explained in detail.

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. A short overview on recent 100 Gbit/s transmission experiments on the basis of a purely electronic receiver and a summary conclude the paper.

2. CARRIER-GRADE ETHERNET-BASED CORE NETWORKS

Next to IP services, VPN business services are generating more and more traffic and revenue for network providers. In particular, Ethernet services (E-Line and E-LAN) are evolving. Today these layer 2 services are mostly transported via IP/MPLS tunnels [1]. However, the complex functionalities and protocols of layer 3 are often not required to transport these pure layer 2 services. Native End-to-end Ethernet structures will arise where Ethernet business services will be transported on pure layer 2 infrastructures without the need of complex data transformations and changes in the functional layer structure.

IP traffic itself still has to be transported by future networks. Routes towards remote IP destinations are learned via BGP

*andreas.kirstaedter@siemens.com; phone +49 89 636 47484; fax +49 89 636 51115; <http://www.siemens.com>

and are distributed between edge routers (or BGP route reflectors) of a network. The appropriate inter-domain route and with it the egress edge router of the network is determined at the ingress edge-router of the network. Following this, today interior gateway protocols are used to define the best path between these edge routers. However, since the source and egress edge routers are well defined, in principle no IP routing functionality needs to be performed inside a core network. Instead, a switching of packets between the two edge-routers is sufficient to transport traffic. More and more MPLS-based approaches are therefore deployed in core networks to build tunnels between all edge-routers. Instead of using MPLS in combination with an Interior Gateway Protocol (e.g. OSPF [2]), end-to-end Ethernet might well be used for this purpose as soon as it offers the correct feature set.

Next to IP and Ethernet services, the transport of high bit-rate data between remote locations can be accomplished at very low cost with WDM technology. Thus, the interactions between Ethernet and WDM have to be considered very closely to achieve optimized multi-layer networks taking into account physical effects, routing, and wavelengths constraints as well as the improvements in cost positions enabled by multi-layer grooming.

In order to be suited for core networks, Ethernet needs carrier-grade performance and functionality. It has to offer and implement the required Quality-of-Service (QoS) and has to enable traffic engineering (TE) to fine-tune the network flows. Furthermore, it has to provide fast and efficient resilience mechanisms to recover from network element failures and has to enable various Operation, Administration, and Maintenance (OAM) features for the configuration and monitoring of the network. Additionally, a high degree of scalability is needed for handling different traffic types and for user separation inside the network. Last but not least it has to provide security of the network. Especially, the scalability in terms of address space, maximum transmission speed, and maximum transmission distance becomes an important issue for the next Ethernet generation. E.g., multi-layer operation and optimization can only be used if facilitated by reasonable values of the maximum transmission distance.

These challenges and the mapping of the functionality of different protocol entities to the requirements of backbone networks is shown in Figure 1 for several possible methods for provisioning path-oriented forwarding. In the following, the main mechanisms that enable carrier-grade Ethernet networks are introduced. It has to be emphasized that the same requirements are often solved in different layers. Thus, a cost-efficient network and protocol architecture has to scrutinize these redundancies between the layers very carefully.

	Scalability	QoS / TE	Resilience	OAM Features
IP	Subnetting, NAT	DSCP	Rerouting	Ping etc.
1) Ethernet: VLAN switching	12 / 24 bit label space	GMPLS-based TE, 802.1p prioritization	1:1 protection switching	IEEE 802.1ag
2) Ethernet: PBT (MAC in MAC)	46 ... (46+12) bit Label space	GMPLS-based TE, 802.1p prioritization	1:1 protection switching	IEEE 802.1ag
3) Ethernet: T-MPLS	20 bit Label space	GMPLS-based TE, Prioritization with EXP bits	1:1 protection switching, FastReRoute	IEEE 802.1ag, LSP ping / LSP traceroute
Ethernet PHY	100Gbps			
WDM	DWDM		standby channels	
Fiber	bundles		redundant fibers	

Figure 1: Protocol layer functionality in Ethernet core networks.

1. Forwarding Technology and Scalability

One of the key issues of End-end-Ethernet networks is scalability. Based on configurable IDs of switches, configurable port weights, and priorities, the Spanning Tree Protocol (STP) [3] calculates one tree-structure to connect any switch with each other. Although loop-less forwarding is guaranteed with this mechanism, STP provides only one path between two locations and a MAC address learning of any equipment is performed at the switches.

When combining large networks and adding hundreds of customer networks with an Ethernet-based core network the number of MAC addresses grows rapidly. Thus, scalability cannot be provided and a separation of networks or an additional hierarchy between them has to be introduced to allow a scalable forwarding of data (Figure 2).

Also, the use of only one tree structure and with it the possibility to use only one path between two locations hamper the use of efficient traffic engineering and resilience mechanisms. Thus, today three connection-oriented forwarding technologies are currently discussed at standardization bodies for Carrier-Grade Ethernet transport networks: VLAN Cross-Connect (VLAN-XC), Provider Backbone Transport (PBT), and Transport Multi-Protocol Label Switching (T-MPLS).

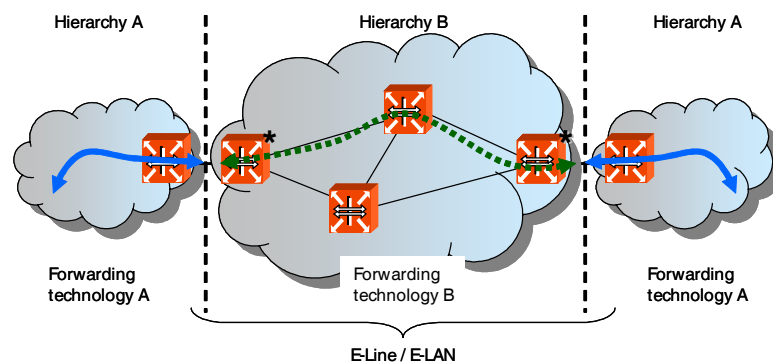


Figure 2: Separation of networks and introduction of a hierarchy.

VLAN Cross-Connect (VLAN-XC)

The main idea of VLAN-XC [4] is to establish pre-defined tunnels between edge switches of a network and to use these tunnels to route and differentiate traffic from each other. Instead of using a destination MAC address for the forwarding decision, a label (VLAN-XC Tag) is encoded in the Ethernet header to determine the appropriate tunnel. Ingress edge-switches have to analyze incoming packets, choose one of the pre-defined tunnels, and label an Ethernet packet

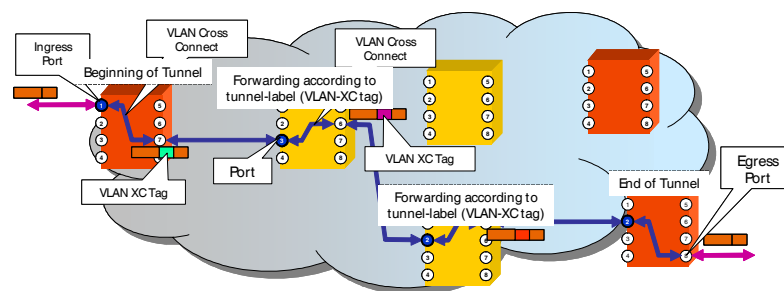


Figure 3: Forwarding of packets using VLAN-XC.

accordingly (see Figure 3). Intermediate switches route the traffic according to the given tunnel label and are able to swap the label. Finally, the tunnel label is removed at an egress switch to allow the transparent transportation of customer data. With this functionality, multiple paths between two edge-switches are supported. Traffic can be separated and distributed in the network and traffic engineering is facilitated.

To avoid changing the Ethernet header structure, VLAN-XC uses the bits reserved for VLAN-IDs of IEEE 802.1Q [5] and IEEE 802.1ad [6] to encode the tunnels. The currently published standardization documents differentiate between two modes: Single tag and double tag forwarding. Figure 4 depicts the VLAN-XC frame structure. The VLAN-XC tags

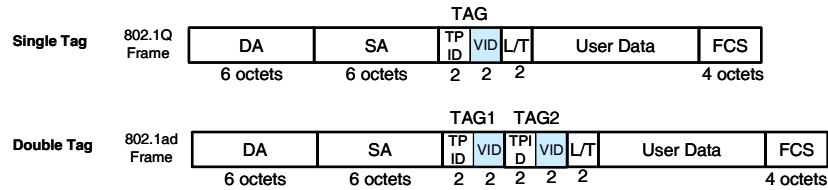


Figure 4: VLAN-XC packet structure using single or double tagging.

are highlighted in blue. For single tagging one tag (12 bits) can be used whereas in double tagging two tags (24 bits) can be combined and used to encode the tunnel.

Provider Backbone Transport (PBT)

Similar to the VLAN-XC, Provider Backbone Transport establishes pre-defined tunnels between edge switches. However, instead of adding a label to the header, a MAC encapsulation is performed at the edge switches (Figure 5).

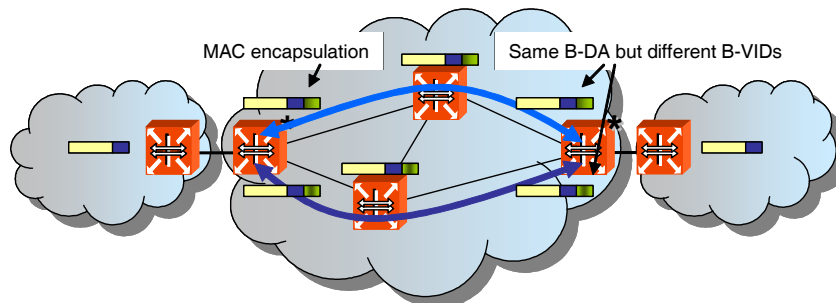


Figure 5: MAC encapsulation in the Core network in PBT.

Figure 6 depicts the frame structure of IEEE 802.1ah [7] used for PBT. The tunnel is encoded by the destination MAC address of the backbone egress switch (B-DA) as well as a 12 bit VLAN-tag (backbone tag, B-VID). PBT forms a topology of B-DA rooted trees and an independent sink-tree is configured for each <B-DA, B-VID> pair. Since no SPT algorithm has to be performed, the trees need not be spanning. Thus, up to 4096 different trees can be configured for one B-DA.

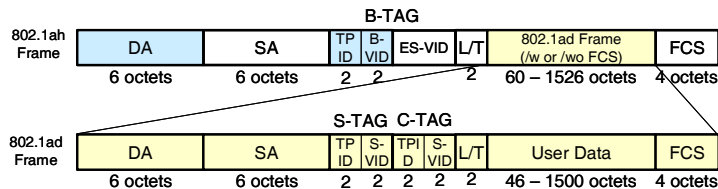


Figure 6: Provider Backbone transport frame structure (MAC in MAC tunneling).

Transport Multi Protocol Label Switching (T-MPLS)

Transport Multi-Protocol Label Switching (T-MPLS) is an adaptation of MPLS and is defined in [8]. The main idea is to use the well established MPLS concept known from IP routing and adapt it for transport forwarding issues. As with VLAN-XC and PBT, T-MPLS establishes pre-defined tunnels. In T-MPLS an additional MPLS header is pushed in front of the client traffic that is transported transparently inside the backbone network. Similarly to VLAN-XC the 20bit label is used to encode the backbone tunnel and is removed at the egress backbone switch. Figure 7 illustrates the frame structure of T-MPLS.

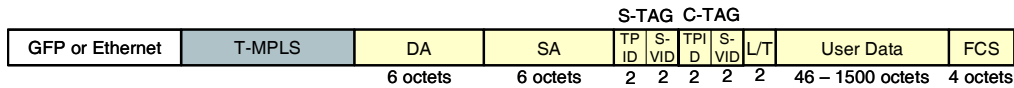


Figure 7: T-MPLS frame structure.

Since there are only limited changes to MPLS we summarize the differences only:

- Separation from the IP control plane, i.e. T-MPLS operates independently of its clients and its associated control networks (management and signaling network).
- Use of Penultimate Hop Popping is prohibited
- Uni-directional and bi-directional Label Switched Paths can be defined
- Use of global or per interface label space. I.e. the label can be swapped at intermediate routers (local port meaning as with VLAN-XC).
- Three types of signaling communication channels (in-band via native IP packets, in-band via dedicated LSP, out-of band)
- The merging of tunnels as well as the equal distribution of traffic onto paths (ECMP) is prohibited.

Scalability is provided by all three proposed forwarding technologies via the introduction of a backbone-network hierarchy for the forwarding of traffic. Edge switches manipulate the incoming Ethernet packets and add tunnel information. Instead of MAC learning (which is disabled inside the core in all the technologies) the forwarding is performed along pre-defined tunnels. The number of tunnels that have to be provided depend on the number of edge-switches, supported types of services, and the number of distinguishable networks (VLANs).

VLAN-XC uses the single or double tagged Ethernet frame to distinguish the tunnels. Since the tags can be swapped at intermediate nodes either $2^{12} = 4096$ or $2^{24} = 16.777.216$ labels or paths can be distinguished per Ethernet port. Single tagging seems to be sufficient in small networks only. However, double tagging with label swapping seems to be sufficient from a scalability point of view.

PBT uses destination MACs in combination with one VLAN-Tag to code the tunnel. The VLAN-Tag in PBT, however, cannot be swapped at intermediate nodes. Thus, $2^{12} = 4096$ different labels or tunnels can be established for each destination address which is sufficient to separate traffic from each other.

T-MPLS pushes an additional header structure to the packet. The MPLS label is swappable and provides $2^{20} = 1.048.576$ different labels. Similarly to VLAN-XC the MPLS label can be swapped at intermediate nodes. Thus, a label has local port meaning only and enough tunnels can be differentiated in the network.

Furthermore, if label-space becomes too small in very large networks all technologies can be extended by introducing an additional hierarchy.

2. End-to-End Quality of Service

Quality of Service (QoS) functionality enables service providers to guarantee and enforce transmission quality parameters (e.g. bandwidth, jitter, delay, packet loss ratio) according to a specified service-level agreement (SLA) with the customer.

A QoS framework that is currently developed by the Metro Ethernet Forum (MEF) aims at providing hard QoS in Ethernet networks [9]. This framework uses the RSVP-TE protocol to setup end-to-end paths with dedicated bandwidth. In native Ethernet networks, traffic is labeled with Service-VLAN tags that are related to a set of QoS parameters. QoS-conform forwarding in Ethernet switches is controlled by GMPLS [10]. In T-MPLS-Ethernet networks, MPLS packets are labeled with T-MPLS tags and forwarded along the specific Label Switched Paths. A connection acceptance control, which is also operated by GMPLS, guarantees that the required bandwidth is available along the requested path.

The MEF's definition of generic service-level parameters enables a high flexibility in SLA definitions: The Committed Information Rate (CIR) determines the minimum amount of bandwidth available to the customer, while the Excess Information Rate (EIR) provides additional bandwidth during low network load periods. Maximum burst sizes corresponding to the CIR and EIR are defined accordingly. For more details on enabling hard QoS in Ethernet networks, please consider [9], [10], and [11].

3. Resilience mechanisms

The usage of the Rapid Spanning Tree Protocol (RSTP) for Layer 2 loop prevention is highly inefficient, since it does not allow packet forwarding along the shortest paths. This means wasting a lot of bandwidth, especially since between any two edge points of a packet core network typically a large amount of bandwidth is needed. Even setting up multiple spanning trees via MSTP (Multiple Spanning Tree Protocol) does not improve the situation: the number of spanning trees which can be handled is usually quite small and creating lots of them would put a big burden on the control unit of the Ethernet switches. Additionally, from a failure restoration point of view RSTP/MSTP is no good choice as it is known to not scale well to large networks.

All three forwarding concepts mentioned in the sections before (i.e. VLAN-XC, PBT and T-MPLS) do not require any MAC learning and no loop prevention mechanism either. In fact, the paths are setup end-to-end either via manual configuration through an appropriate network management system or using a dynamic control plane like GMPLS. The use of GMPLS as an Ethernet Control plane is called GELS (GMPLS controlled Ethernet Label Switching). Further investigations in this area are done within the IETF working group CCAMP. Additionally, the MEF protection framework [13] presents additional, generic resilience concepts that aim at enabling interoperation between future Ethernet devices of different vendors.

For failure protection, the concepts known from MPLS can be used. Working as well as backup paths can be pre-established (protection) or can be calculated and set-up if required (restoration). Thus, for protection, backup paths can be pre-established for considered failure patterns that protect each path or a group of paths while also taking into account the required bandwidth and QoS constraints.. When a failure occurs, the ingress node will be informed about the failure almost immediately either by signaling through RSVP-TE or by some Hello mechanism on the paths (similar to the concept of Bidirectional Forwarding Detection – BFD). When the ingress node knows about the failure, he can initiate a switchover to the backup paths very rapidly. Protection mechanisms like Fast Reroute might also be used for an extremely fast local repair to optimize the failover time even more.

Using appropriate mechanisms, a fast failover can be guaranteed for all failures which happen between any two Ethernet devices within the core network. Other failures, like a link cut between an edge router and an Ethernet device cannot be taken care of by the Layer 2 core itself but the need to be detected and handled by the edge routers themselves.. It is well known that the IGP can be tuned to achieve acceptable rerouting failover times in such a situation.

Altogether, depending on the failure scenarios (single or multiple failures) a network provider wants to protect his infrastructure against, the resilience mechanisms in all layers can be combined appropriately. Especially when protection schemes are used throughout several layers, the configuration has to make sure that the protection/restoration mechanisms do not badly influence each other. This is usually realized by assuring that in case a failure is detected in a certain layer which also can be repaired within that layer, no higher layer should even recognize the failure. This leads to a well designed multi-layer resilience scheme.

4. Operations Administration & Maintenance

Enhanced OAM functionality is indispensable in carrier networks as it enables failure detection, localization, and performance monitoring. OAM features include Fault Management, Configuration Management, Accounting / Administration, Performance Management, and Security Management (FCAPS).

OAM functionalities for Carrier Grade Ethernet are currently in standardization [14] [15]. These functionalities include auto-detection, connectivity checks, fault detection and alarming, link traces, and loopback for the facilitation of configuration and failure detection as well as procedures for the measurement of different QoS performance parameters/features for point-to-point connections (e.g. frame loss ratio, delay, jitter, and throughput). Additionally, security management functionalities like the detection of denial of service attacks as well as resource and admission control procedures in combination with (dynamic) re-optimization of the network paths can be provided.

3. MULTI-LAYER OPTIMIZATION

Another important aspect in the area of scalability is the maximum transmission distance of Ethernet signals. Multilayer network approaches are very attractive for the purpose of reducing unnecessary packet processing in intermediate nodes

[16]. Via multi-layer grooming transit traffic can bypass intermediate nodes as shown in Figure 8. Traffic between two edge switches can either be transported transparently in the optical domain or can be converted to the electrical domain to allow grooming along the path.

As in backbone networks node distances are usually in the area of a few hundred kilometers or even more the maximum transmission distance has to be considered carefully at multi-layer approaches. E.g., the maximum transmission distance of current 10G Ethernet is to 70-80km according to vendor specifications, which is already way above the officially declared maximum range of 40km.

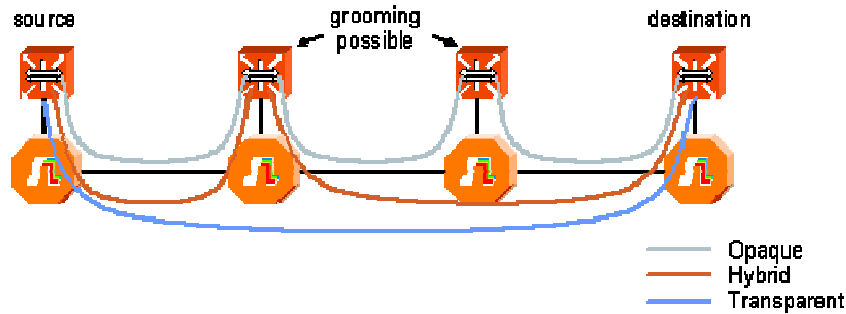


Figure 8: Multi-layer routing.

When speeding Ethernet up to 100 Gbit/s, the transmission over long distances will become even more difficult: Second degree (slope) chromatic dispersion has to be exactly compensated, birefringency effects become grave, and the signal-to-noise ratio of 100 Gbit/s signals is generally lower as fewer photons are transmitted per optical impulse. Although several experiments and field trials of the past (e.g. that of Siemens, British Telecom, and the University of Eindhoven in the EU IST project FASHION with 160 Gbit/s signals over 275km standard single-mode fiber [17]) showed that long-distance transmissions of high-speed signals are definitely feasible, the complexity and costs of the required optical equipment is still at a very high level.

On the other hand, the effort spent on extending the signal reach of Ethernet signals is rewarded by equipment savings. Figure 9 illustrates 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. These results were derived for an Ethernet network following the topology of a generic German backbone (Figure 9) and are also used during the CAPEX and OPEX analyses below [18].

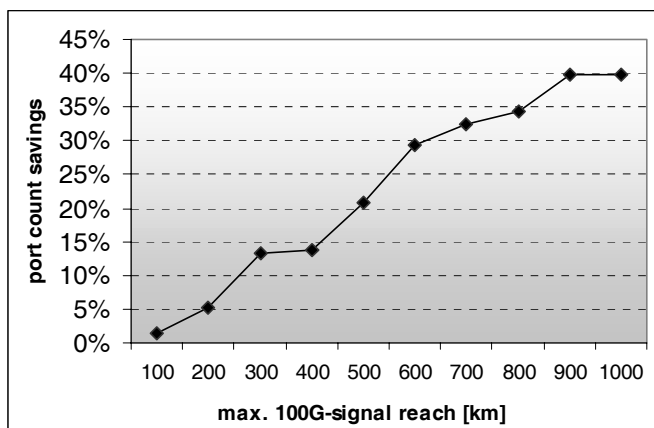


Figure 9. Port-count savings in grooming-enabled Ethernet networks.



Reference network topology.

IEEE 802.3 standards usually define a broad set of physical Ethernet interfaces and besides a single 100G-Ethernet signal, a multiplexing of lower bit-rate optical signals into a 100G signal are also conceivable (e.g. 10x10G or 4x25G).

Although the maximum transmission range of these multiplexed signals would certainly be longer than that of a pure 100G-signal, the multiplexing requires additional WDM equipment and the signal would occupy several wavelength channels on the fiber.

4. CAPEX AND OPEX PERFORMANCE OF ETHERNET CORE NETWORKS

1. CAPEX: Modeling Approach and Results

In order to calculate the total CAPEX of 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. Corresponding traffic matrices were extrapolated to determine future link loads. The traffic is assumed to grow homogeneously at a rate of 40% per annum. A shortest-path routing algorithm has been 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 fiber 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 [19]).

2. Architectures and Results

The following *generic* network architectures were considered:

(a) IP/POS-over-WDM: The backbone consists of Label Edge Routers (LERs) and Label Switch Routers (LSRs) that are all equipped with POS interfaces. SONET/SDH is only used for transporting the IP packets node to node, directly over WDM. A 1+1 protection scheme is applied.

(b) IP/POS-over-SDH-over-WDM: Similar to (a), this network scenario considers a backbone where LERs are located at the ingress and egress points of the backbone. However, the traffic is switched and groomed along SDH add-drop-multiplexers and cross-connects inside the core. Thus, no LSRs are required inside the network. A 1+1 protection scheme is applied.

(c) IP/POS-over-OXC-over-WDM: This case is similar to the previous architecture. However, the SDH switches are replaced by optical cross-connects (OXCs). The range of the optical signal is assumed to be large enough to enable end-to-end optical grooming. Thus, again no LSRs are required inside the network. A 1+1 protection scheme is applied.

(d) IP/MPLS-over-Ethernet-over-WDM: The backbone consists of LERs at the edge and MPLS-enabled Ethernet switches in the core of the backbone. This architecture corresponds to MPLS Ethernet architecture as outlined in section 2. A 1:1 protection scheme is applied, i.e. all capacity is over-provisioned by 100%.

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

(f) Ethernet-over-WDM with service-level protection: The network architecture is identical to (e) except that only premium traffic is protected against failure (share of premium traffic set to 30%).

The architecture-dependent CAPEX components (aggregated for 2009 to 2012) are illustrated in Figure 10 - split up into components belonging to the different layers. The high prices of POS interfaces generally lead to a high cost component of LERs for all SDH-related infrastructures. 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 100 Gbit/s Ethernet network enables higher savings in the LER category. Applying a service-level differentiated protection scheme, the CAPEX can be reduced even further.

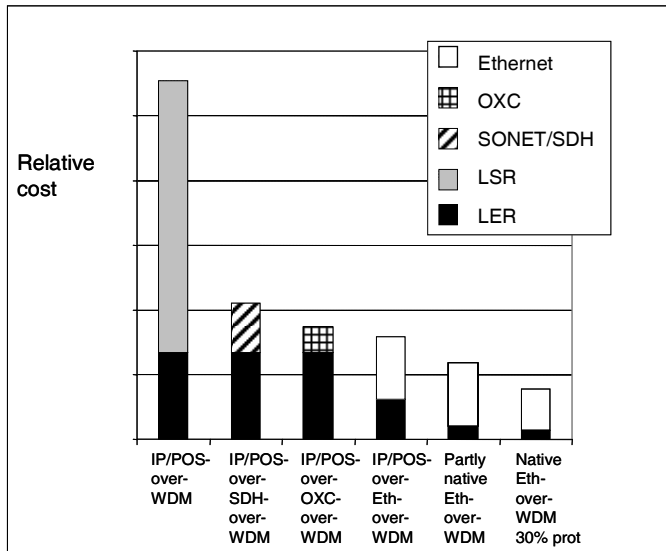


Figure 10. Accumulated CAPEX comparison.

3. OPEX Comparison

Generally, OPEX are evaluated via process-oriented approaches [20]. While OPEX include a lot more processes, the repair process is selected here in this paper since the impact of 100 Gbit/s Ethernet gets most visibly in this area. *Note:* WDM failures and fiber breaks again are not considered as they are common for all different architectures. Thus, the total OPEX repair process values are very low.

The OPEX were evaluated for each network architecture (and year) described above via first determining the total number and type of equipment. By using availability figures [21] 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 1 shows the results of the repair process OPEX, accumulated over the years 2009 to 2012 (OPEX rising with network growth):

Table 1. OPEX repair-process comparison, accumulated.

NW architecture	Normalized cost
(a) POS over WDM	3.0
(b) POS over SDH	2.0
(c) POS over OXC	2.8
(d) IP/MPLS over 100GE	1.4
(e) Native 100GE	1.4
(f) Native 100GE, 30% protected	1.0

Due to the reduced device count (less switches and line cards) enabled by 100 Gbit/s Ethernet instead of 40Gbit/s POS Ethernet networks are more economical than SDH-architectures. A service-level protection scheme further reduces the required network transport capacity, the network element count, and thus 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 (see study conducted by the MEF [22]). Thus, Ethernet architectures enable considerable savings in a wide range of OPEX areas.

5. 100 GBIT/S TRANSMISSION EXPERIMENTS

The sections above showed that Ethernet operation at speeds of 100 Gbit/s is very desirable in terms of architecture-related network cost. The transmission of high speed data rates above 100 Gbit/s is well understood and can be managed. As a consequence, the knowledge to realize the transmission of a 100 Gbit/s Ethernet signal is present. The problem still to solve is to find efficient electro-optical and opto-electrical conversion techniques. Electrical solutions are preferable to handle the data at the transmitter and receiver since OTDM techniques are still too complex and difficult to implement in commercial products.

By using ultra-fast electronic circuits instead of elaborate optical methods in high-capacity optical transmission systems cost per transmitted bit per second and kilometer can be reduced. Electronic circuitry for 40 Gbit/s is already commercially available. To really exploit the cost advantage of an electrical receiver compared to optical solutions a compact integrated device is needed - preferably a single chip.

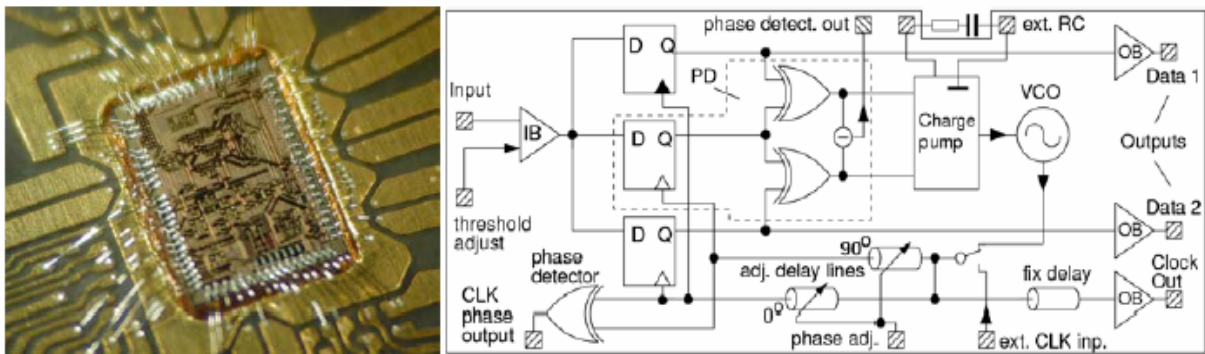


Figure 11. Photo (left) and block diagram (right) of the integrated ETDM receiver chip [23].

Recently, as an important step towards 100 Gbit/s Ethernet an integrated ETDM receiver comprising 1:2-demultiplexing (DEMUX) and clock & data recovery (CDR) on a single chip was presented [23]. This receiver was tested in a 100 Gbit/s transmission experiment. Error-free performance ($BER < 10^{-9}$) was obtained back-to-back and after transmission over 480 km of dispersion managed fiber. The ETDM receiver was initially designed for 80 Gbit/s operations. A redesign of the receiver chip is expected to enable an even better performance and operation at even higher bit rates.

6. SUMMARY AND CONCLUSIONS

In the past, Ethernet evolved from LAN into Metro areas covering speeds from 10 Mbit/s up to 10 Gbit/s. Next-generation Ethernet transmission-speeds of 100 Gbit/s will be the enabler of Ethernet-based pure packet core networks and facilitate a cost-efficient Ethernet transport. Carrier-gradeness of Ethernet-based packet architectures is the major challenge. A careful analysis of the required protocol features like network resilience, QoS, and OAM shows many approaches for solving the hierarchy and scalability problems of backbone networks. However, redundancies between the layers of today's network architectures have to be resolved in order to shape a new end-to-end Ethernet layer with the required scalability.

A CAPEX and OPEX analysis demonstrates a considerable cost advantage of 100 Gbit/s carrier-grade Ethernet in comparison to IP/MPLS/SDH-based solutions. Native Ethernet networks promise the highest savings in terms of OPEX and CAPEX. 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 100G-Ethernet networks and the efficient economics of Ethernet services are responsible for a superior OPEX performance.

Therefore, it can be said that 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.

ACKNOWLEDGEMENTS

We would like to thank A. Schmid-Egger for the input to the CAPEX and OPEX analysis. Special thanks go to the whole E3 project team at Siemens AG, namely R. Derksen, G. Lehmann, S. Pasqualini, T. Fischer, M. Hoffmann, G. Göger, and D. Schupke. We would also like to thank the colleagues from Heinrich-Hertz-Institute in Berlin, namely C. Schubert for all valuable discussions and all contributions to the measurement results of the 100 Gbit/s transmission experiment, and M. Möller from Micram Microelectronics for the support upgrading the integrated receiver from 86 to 100 Gbit/s. The financial support of the Bundesministerium für Bildung und Forschung (BMBF) is gratefully acknowledged.

REFERENCES

1. E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture," IETF Request for Comments 3031, January 2001, [Online] Available: <http://www.ietf.org>
2. J. Moy, "OSPF Version 2," IETF Request for Comments 2327, April 1998, [Online] Available: <http://www.ietf.org>
3. ANSI/IEEE Standard 802.1D, "MAC Bridges"
4. D. Papadimitriou, N. Sprecher, et al., "A Framework for GMPLS-controlled Ethernet Label Switching," IETF Draft, February 2006, draft-dimitri-gels-framework-00.txt, [Online] Available: <http://www.ietf.org>
5. ANSI/IEEE Standard 802.1Q, "Virtual Bridged Local Area Networks"
6. ANSI/IEEE Standard 802.1ad, "Provider Bridges"
7. ANSI/IEEE Standard 802.1ad, "Provider Backbone Bridges"
8. ITU-T Recommendation G.8110, "T-MPLS Layer Network Architecture," January 2005
9. 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
10. 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>
11. MEF, "Bandwidth Profiles for Ethernet Services" [Online]. Available: <http://www.metroethernetforum.org/PDFs/WhitePapers/Bandwidth-Profiles-for-Ethernet-Services.pdf>
12. Atrica Corp., "Delivering Hard QoS in Carrier Ethernet Networks," [Online]. Available: http://www.atrica.com/body/products/whitepapers/Delivering_Hard_QoS_in_Carrier_Ethernet_Networks_v12.pdf
13. MEF, "Requirements and Framework for Ethernet Service Protection in Metro Ethernet Networks", [Online]. Available: <http://www.metroethernetforum.org/PDFs/Standards/MEF2.pdf>
14. ITU-T Recommendation Y.1730 – Requirements for OAM functions in Ethernet based networks, January 2004
15. ITU-T Recommendation Y.1731 – OAM Functions and Mechanisms for Ethernet based Networks, January 2006
16. M. Scheffel, R. G. Prinz, C. G. Gruber, A. Autenrieth, D. A. Schupke, „Optimal Routing and Grooming for Multilayer Networks with Transponders and Muxponders“, IEEE Globecom, San Francisco, USA, 2006.
17. J.P. Turkiewicz, E. Tangdionga, 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. D. A. Schupke, M. Jäger and R. Hülsermann, "Comparison of Resilience Mechanisms for Dynamic Services in Intelligent Optical Networks", Fourth International Workshop on the Design of Reliable Communication Networks (DRCN), Banff, Canada, 2003.
19. A. Schmid-Egger, A. Kirstädter, "Ethernet in Core Networks:: A Technical and Economical Analysis", Proc. HPSR2006, Poznan, Poland, June 2006.
20. S. Pasqualini, A. Kirstädter, A. Iselt, R. Chahine, S. Verbrugge, D. Colle, M. Pickavet, P. Demeester, "Influence of GMPLS on Network Providers' Operational Expenditures: A Quantitative Study", IEEE Communications Magazine, Vol. 43, No. 7, pp. 28-34, July 2005
21. NOBEL Project, "Availability Model and OPEX related parameters", to be published .
22. MEF, "Service Provider Study: Operating Expenditures," [Online]. Available for MEF members: <http://www.metroethernetforum.org>
23. 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 Transmission Experiment over 480 km DMF“, Proc. of OFC 2006, PDP37.